An investigation into the relationship between adolescent risk taking behaviour and executive functions, impulsivity and inner speech.

Hyde, Zoe Harriet

Awarding institution:
King's College London

VOLUME I

Systematic Literature Review

Empirical Project

Service Evaluation Project

Zoe Hyde

Institute of Psychiatry, Psychology and Neuroscience

King's College London

Thesis submitted in partial fulfilment for the degree of

Doctorate in Clinical Psychology 2015

Acknowledgements

I would like to thank my first supervisor Troy for all of the interesting conversations, the insightful supervision and for keeping me calm and laughing. Essi and Eamon's critiques and advice, and Daniel's expert statistical knowledge, were also very much appreciated. I would like to thank all of the teachers and the pupils who gave their time to make this research project possible, especially Marilyn and Cath. I am very grateful to Ben for his expertise in the field of inner speech and to Suvena for giving her time and support. I would like to thank Amita for her valuable supervision of the service project. I would like to acknowledge the clients that the case studies are based on and thank my supervisors for their support and feedback, especially Charlotte, who slogged through two of them! Finally I would like to say thank you to my brilliant parents, for their never-ending encouragement, patience and care, and to my wonderful cohort, without whom this would have been so much harder.

Overall Table of Contents

Systematic Literature Review

A review of the methods used to investigate individual differences in inner speech and whether these individual differences are related to psychopathology

Supervisor: Dr Troy Tranah.

Discussant: Dr Ben Alderson-Day

Co-rater: Dr Suvena Sethi.

Table of Contents

List of Figures and Tables

Abstract

This systematic literature review critically appraised and compared methods for measuring individual differences in inner speech in developmentally neurotypical adults. It also investigated whether individual differences in inner speech which were found through these methods of measurement, were related to psychopathology. Literature searches were performed in three databases. Twenty nine studies were included in this review.

These studies measured inner speech through a variety of methods including questionnaire and interview, introspective, experimental, neuroimaging and lesion methodologies. Methods varied with regard to the characteristics of inner speech they measured, as well as the qualities of the inner speech which the measurement method relied on. Only two of the included studies used two different types of inner speech measurement. This combining of different methods enabled reciprocal validation of each method, and indicated that some findings may be artefacts of the particular method used.

Individual differences in inner speech found by these methods of measurement, were related to variation in other aspects of human experience. This included variation in psychological disorders and symptoms such as auditory verbal hallucinations, anxiety and self-esteem. These preliminary findings suggest that a more thorough assessment of the relationship between inner speech processes and psychological disorders and symptoms would be useful for the advancement of clinical psychology theory and treatment.

Introduction

Inner speech

Inner speech has been defined as "the subjective experience of language in the absence of articulation" (Alderson-Day and Fernyhough, in press, p. 1). This phenomenon is considered to be an important mental activity, which humans engage in every day (Geva, Bennett, Warburton, & Patterson, 2011), and it has been suggested that "most people hear a little voice inside their head when thinking, reading, writing and remembering" (Oppenheim & Dell, 2008, p. 529). Socrates defined thought, "as the talk which the soul has with itself … the soul … when it thinks, is merely conversing with itself, asking itself questions and answering" (Plato, 1987, 189e) (McCarthy-Jones & Fernyhough, 2011, p. 1586), suggesting that this phenomenom has been a long-standing aspect of the human experience.

Hurlburt and colleagues (2013) found that people tend to perceive inner speech (or 'inner speaking' as they refer to it) as an experience that is similar to speaking aloud, with the only difference between these two phenomena being people's "immediate and unshakeable recognition that the speaking is inner rather than external" (Hurlburt, Heavey, & Kelsey, 2013, p. 1482). These authors report that people tend to experience themselves as producing inner speech, rather than it being something which they are hearing. That is, people tend to feel that the words are under their control, and are not being passively received through an auditory format (Hurlburt et al., 2013).

Development of inner speech

Vygotsky was the first researcher to suggest that inner speech is the outcome of a developmental process (Vygotsky, 1962). In his theory of cognitive development, Vygotsky (1962) hypothesised that children initially depend on external instruction, from their caregiver for example, to guide their behaviour. They gradually learn to do this independently, by using an outer monologue, termed 'private speech', which is eventually internalised as inner speech (Vygotsky, 1962).

Winsler and Naglieri (2003) conducted a large cross-sectional study of 2,156 children and adolescents aged from 5 to 17 years old. They observed that although 43% of 5 year olds used private speech while completing a planning task, this declined linearly with age and only

10% of 17 year olds were found to engage in overt self-talk. In contrast, self-reported engagement in inner speech increased with age, and while only 4% of the youngest children reported using inner speech, 30% of the oldest adolescents did so (Winsler & Naglieri, 2003). The frequency of partially covert whispers and mutterings was observed to follow an inverted 'U', being apparent in 13% of 5 year olds, peaking at 28% for 9 year-olds, and decreasing to 11% in teenagers (Winsler & Naglieri, 2003). It appears that overt private speech is common in children during their preschool years, peaking at around 5 years of age and becoming less frequent in children aged 5 to 8 (Winsler, 2009). This decline is accompanied by less audible private speech (Berk, 1986) and parallels the growth of self-reported inner speech use (Damianova, Lucas, & Sullivan, 2012). These studies support Vygotsky's (1962) theory, that once-external self-directed verbalisations gradually become internal.

Vygotsky suggested that the process of internalisation during normal development involves syntactic and semantic changes that result in fully internalised inner speech having little resemblance to the outer dialogue from which it originated (Fernyhough, 2004; Vygotsky, 1962). Fernyhough (2004) elaborated on these ideas to suggest that four different levels of internalisation occur during the development of inner speech. In this model Level 1 refers to full externalised dialogue with other people. Level 2 refers to 'private speech' where external dialogue is directed just to oneself. Level 3 is termed 'expanded inner speech', and is understood to be silent speech which still has the syntactic and semantic qualities of overt speech. Finally, Level 4 is termed 'condensed inner speech', and refers to inner speech which has undergone syntactic and semantic changes (Fernyhough, 2004). In comparison with the trajectories for the transition from private to silent self-talk, the developmental timings of the transition between Level's 3 and 4 in Fernyhough's (2004) model have not yet been fully delineated.

Fernyhough (2004) suggests that individuals can move between the levels of self-directed talk in his model throughout their life. That is, while adults engage in condensed speech by default, more demanding cognitive conditions will lead an individual to engage in expanded inner speech, and particularly stressful situations may require someone to use private speech (Fernyhough, 2004). Indeed the phenomenom of private speech has been observed to persist into adulthood (Duncan & Cheyne, 2002), even after individuals have fully internalised inner speech. Duncan and Cheyne (2002) observed that adults engaged in private speech when they were completing a difficult task, but as their competence on the task increased, their

overt self-verbalisations decreased. This finding indicates that adults use private speech as a tool to aid self-regulation, and to support their learning of challenging new skills.

Vygotsky's work was paralleled by that of Luria (1965), who theorised that inner speech could support self-regulation through its interactions with the executive functions. Luria (1965) suggested that children are born with certain prelinguistic capacities that enable monitoring, inhibition and planning of behaviour. Luria (1965) proposed that when these basic capacities enter into relations with language abilities, the executive functions begin to emerge in more developed forms (Luria, 1965; Alderson-Day & Fernyhough, in press). As suggested by Luria's (1965) basic capacities, the executive functions are understood to be processes that allow people to adapt to their environment, through making plans and inhibiting unsuitable behaviour (Jurado & Rosselli, 2007). Anderson et al. (2001) propose that the executive functions encompass three processes, which include 'attentional control', 'cognitive flexibility' and 'goal setting' (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001). The inner speech system is understood to be a separate system, which can be recruited as a cognitive tool to support the executive functions (Alderson-Day & Fernyhough, in press). However inner speech is a linguistic process which can also be active separately from working memory and other executive processes (Geva, Bennett, et al., 2011).

The scientific study of inner speech

Inner speech has been considered difficult to study empirically because it has no obvious behavioural manifestation and so it is difficult to observe (Alderson-Day & Fernyhough, in press). However recent years have seen advances in the development of the scientific study of inner speech, and novel methods of investigation are now being applied to examine this phenomenom (McCarthy-Jones & Fernyhough, 2011; Vicente & Martinez Manrique, 2011).

In addressing the challenge of measuring inner speech, dual-task paradigms are one of the methods that have been used to investigate relationships between inner speech and various cognitive functions. Typically in these investigations, a secondary task is used to disrupt inner speech. For example, articulatory suppression tasks require individuals to say simple distractor words while they complete a primary task which investigates the cognitive process of interest. Research has focused on cognitive flexibility and has found that participants' performance is detrimentally affected by articulatory suppression on tasks where they are

required to switch between different rules, for example (e.g. Baddeley, Chincotta, & Adlam, 2001). Planning ability is also impaired by concurrent articulatory suppression, especially in investigations with children (e.g. Lidstone, Meins, & Fernyhough, 2010).

Such studies provide evidence for some of the functions that inner speech may have. However these dual-task methods investigate the importance of inner speech in relation to other cognitive processes, and not as a stand-alone phenomenon. That is, they do not investigate inner speech as a process in its own right.

Several questionnaires have been developed specifically for measuring self-directed talk. Questionnaires which were validated in adult samples include the Scale for Inner Speech (SIS) (Siegrist, 1995), which investigates the use of self-directed talk for solving personal problems (Schneider, Pospeschill, & Ranger, 2005), the Self-Verbalization Questionnaire (SVQ) (Duncan & Cheyne, 1999), which is designed to measure the use of private speech (Uttl, Morin, & Hamper, 2011), the Self-Talk Use Questionnaire (STUQ) (Hardy, Hall, & Hardy, 2005), which relates to the use of self-talk specifically within sport, and the Self-Talk Scale (STS) (Brinthaupt, Hein, & Kramer, 2009), which can be used to measure the affective tone of self-directed verbalisations (McCarthy-Jones & Fernyhough, 2011). These questionnaires typically explicitly state that they are asking about both externalised self-talk, as well as inner speech, and none of them differentiate between inner and private speech. Uttl et al. (2011) investigated the relationships between these questionnaires in an undergraduate sample and found that all scales had at least acceptable reliabilities (as measured by Cronbach's alpha). However there was very little convergent validity between the findings of these different scales, indicating that these measures investigate different aspects of people's self-directed talk. In contrast with these questionnaires, the Varieties of Inner Speech Questionnaire (VISQ) (McCarthy-Jones & Fernyhough, 2011) specifically measures inner speech and was designed to assess various phenomenological properties of inner speech, such as the extent to which individuals report that their inner speech is condensed or expanded.

Other questionnaires have been developed to measure the frequency with which people think that they are having particular cognitions, often relating to positive or negative content. For example, the Automatic Thoughts Questionnaire (ATQ) (Kendall, Howard, & Hays, 1989) asks people to determine the frequency with which they experience specific, pre-scripted, positive or negative thoughts (e.g. "I'm worthless"). The Self-Talk Inventory (STI) (Burnett,

1996), which was designed primarily for use with children, asks about the content of self-talk by presenting scenarios and asking whether participants think that they would engage in positive (e.g. "everything will be ok") or negative (e.g. "this is going to be awful") overt or covert self-directed statements, which are also pre-scripted. Such methods focus on specific scripted thoughts, which have been determined to be of interest a priori (Heavey & Hurlburt, 2008).

Individual differences in inner speech and variation in other experiences

Contrary to the claims by some researchers that people are always engaging in inner speech (e.g. Baars, 2003), Hurlburt et al. 's (2013) lab have found that some individuals report speaking to themselves silently the majority of the time, while others have almost no experiences of inner speech. In Heavey and Hurlburt's (2008) study, for example, the average frequency of inner speaking across the 30 participants assessed was 26%, while the frequency within participants ranged from 0% to 75%. There is therefore large inter individual variation in these frequencies, and Hulburt's methods have indicated that while some people inner-speak 0% of the time, others are observed to inner-speak nearly always when they are awake (Hurlburt et al., 2013).

Individual differences in the use of self-talk have been associated with individual differences in people's ability to self-regulate. Children's use of private and partially internalised speech during task performance has been related to independent observations of their on-task behaviour and affect in the classroom, as well as to parent and teacher reports of their behavioural difficulties (Winsler, De Leon, Wallace, Carlton, & Willson-Quayle, 2003). In adults, inner speech has been found to support sporting performance (Hatzigeorgiadis, Zourbanos, Galanis, & Theodorakis, 2011). In fact self-talk interventions, in which participants are instructed to engage in self-talk in a particular format, have been related to improvements in athletes' sporting abilities. For example, Chang et al. (2014) found that when participants' engaged in motivational self-talk (e.g. "I believe I can do it well"), they threw a soft-ball further and with more accuracy than when they engaged in self-talk that was unrelated to their sporting behaviour. However, Alderson-Day and Fernyhough (in press) suggest that the self-regulation of cognition and behaviour is just one of the functions of private speech. For example, children's private speech has also been related to creativity (e.g.

Daugherty, White, & Manning, 1994), social skills (Winsler et al., 2003) and emotion regulation (e.g. Day & Smith, 2013).

Inner speech variation and psychopathology

Given these relationships between self-directed talk and various dimensions of human experience, it is possible that individual variation in aspects of inner speech such as frequency (Hurlburt et al., 2013) is related to psychological disorders and symptoms. Such a hypothesis corresponds to research in mental imagery, another form of inner experience, which is suggested to play a key role in various psychological disorders (Pearson, Deeprose, Wallace-Hadrill, Heyes, & Holmes, 2013), such as post-traumatic stress disorder (Holmes, Grey, & Young, 2005) and depression (Patel et al., 2007).

With regard to inner speech, one hypothesis is that the raw material that contributes to the experience of auditory verbal hallucinations, may actually be an individual's own inner speech, which is misattributed to an external source (Jones & Fernyhough, 2007). Evidence is also emerging to suggest that the experience of worrying is linked to verbal thinking, as opposed to mental imagery (Behar, Zuellig, & Borkovec, 2005; Stokes & Hirsch, 2010). For example, Stokes and Hirsch (2010) found that individuals from a non-clinical sample experienced an increase in negative intrusions when instructed to worry in verbal formats, as opposed to engaging in visual imagery which was associated with a decrease in such intrusions. A less well evidenced hypothesis is that verbal thinking has a role in the maintenance of depression (Alderson-Day & Fernyhough, in press). Some support for this claim was provided by Holmes and colleagues (2006) who found that when individuals with depression were instructed to think in verbal form about a possible scenario they experienced a reduction in their mood and an increased susceptibility to a subsequent mood induction, relative to engaging in visual mental imagery.

The current systematic literature review

As far as is known, there has been no previous review of the methods used to assess individual difference in inner speech as a process in its own right. This systematic review aims to critically appraise and compare these different methods. A secondary aim of this review is to investigate whether these methods have found evidence to suggest that inner speech variation is related to psychopathology.

Research question

What methods are used to investigate individual differences in inner speech? Is there a relationship between these individual differences and psychopathology?

Method

This literature review was conducted in accordance with the PRISMA checklist (Moher, Liberati, Tetzlaff, & Altman, 2009).

Data Sources

The electronic databases Medline, PsycInfo and Web of Science were searched in order to obtain the literature. Searches in Medline and PsycInfo were run using OVID search tools, while Web of Science has its own search tool. These searches were conducted on 22nd November 2014.

Search Strategy

Terms relating to 'inner speech' were combined with terms relating to 'measurement' in all three databases. Relevant subject headings were included in Medline and PsycInfo database searches, along with key-words, and truncation was used where possible to capture variation in the terminology. Specifically, the inner speech terms included 'self-talk', 'inner speech', 'internal speech', 'internal dialogue', 'inner dialogue', 'internal monologue', 'inner monologue', 'self-directed speech', 'verbal mediation' and 'private speech'. The measurement terms included 'measur*', 'psychometric*', 'assess*', 'questionnaire*', 'introspect*', 'quantif*', 'examinat*', 'evaluat*', and 'investigat*'. See Appendices 1, 2 and 3 for PsycInfo, Medline and Web of Science search strategies respectively.

In total, 1,976 papers were found using these search criteria in the three databases, of which 643 were duplicates. A total of 1,333 references were included, once duplicates were removed.

Inclusion and Exclusion Criteria

This review included papers that were published in a peer reviewed journal after 1970 and written in English. Papers were included if they empirically investigated individual differences in inner speech in developmentally neurotypical human adults. Lesion studies were therefore included (so long as the lesion had occurred in adulthood), these individuals being assumed to have developed typically.

Non-empirical literature was excluded, among them reviews, solely theoretical articles and books. Papers were also excluded if they looked at the relationship between inner speech and learning a new skill (e.g. learning a new language), or if they investigated the involvement of inner speech in a particular task which was pre-determined by the experimenters (and was therefore supposing a priori that inner speech could influence that particular process). In addition investigations into the impact of interventions on inner speech were excluded. This included studies exploring the role of verbal thinking on rumination and worry, where participants were instructed to think in a particular way (e.g. either verbally or in imagery). These exclusion criteria were applied in order to focus the review on naturally occurring inner speech as a process in its own right.

Papers were also excluded if they investigated how frequently participants had specific thoughts that were pre-defined by experimenters. Endorsing the experience (or not) of scripted, non-idiosyncratic cognitions, was not considered to be a valid form of investigating the naturally occurring phenomenon of individuals' inner speech.

This review focused on adult populations, because the trajectory of inner speech development has not yet been clearly delineated. Papers were therefore excluded if they investigated inner speech in older adult, child or adolescent populations, in order to ensure that the method of assessment was not reporting individual differences resulting from the sample being at different points on the development trajectory of inner speech. Papers that did not clearly explore inner speech as a separate process from private (i.e. overt) speech were excluded. This was to ensure that the review focused on methodologies that can be used to investigate the non-observable phenomenon of inner speech.

The titles of the 1,333 references produced by the electronic database search were screened in accordance with the inclusion and exclusion criteria. The categories of 'eligible', 'query' and

'not eligible' were used. The 207 references that were queried or eligible at title were screened at abstract according to the inclusion/exclusion criteria, using the same three categories. Finally the inclusion/exclusion criteria were applied to the 99 references that were queried or eligible at abstract (see Figure 1).

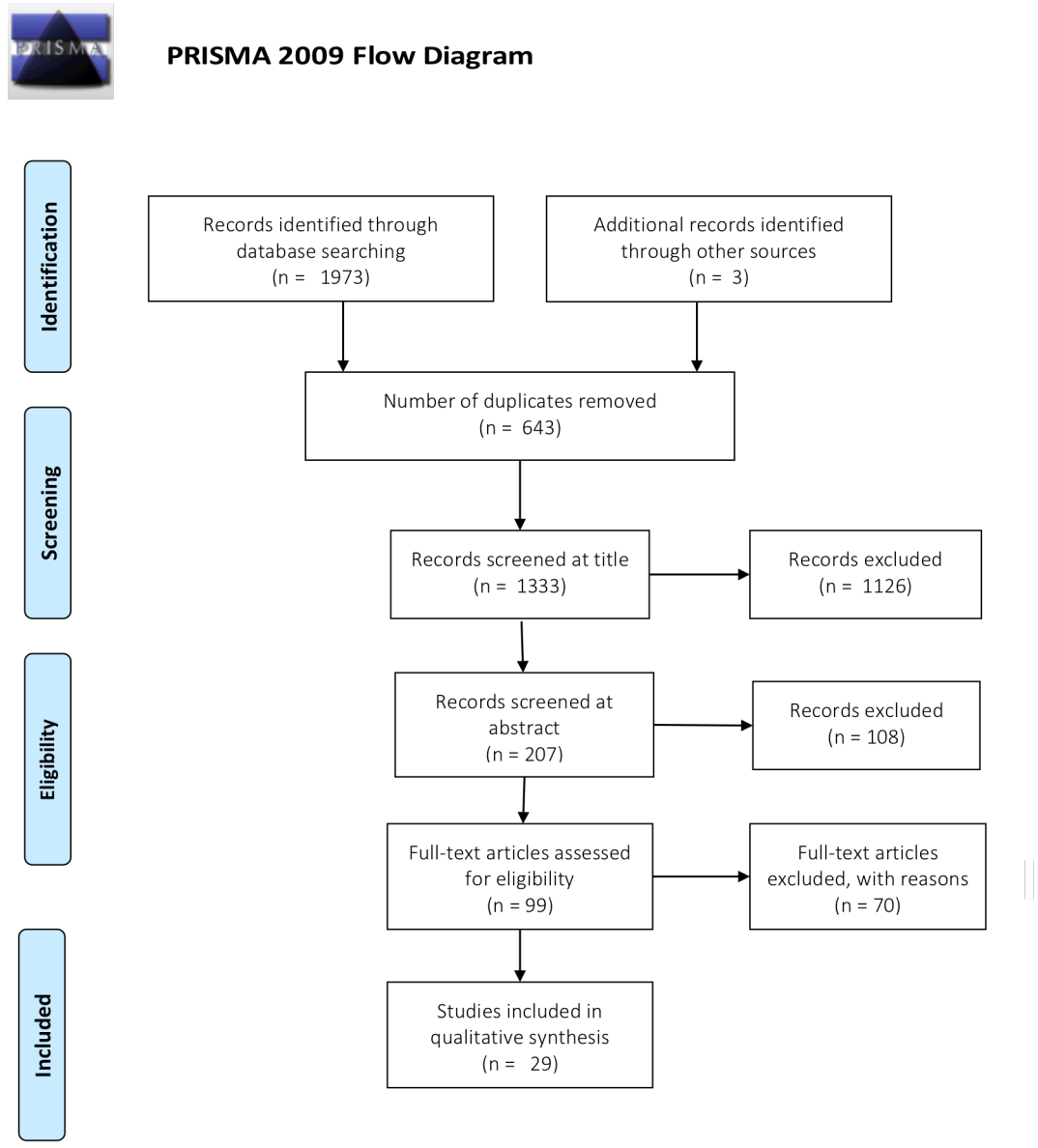


Figure 1_PRISMA map of the number of records identified, included and excluded

Reasons for exclusion were reported for full-text articles (see Figure 2).

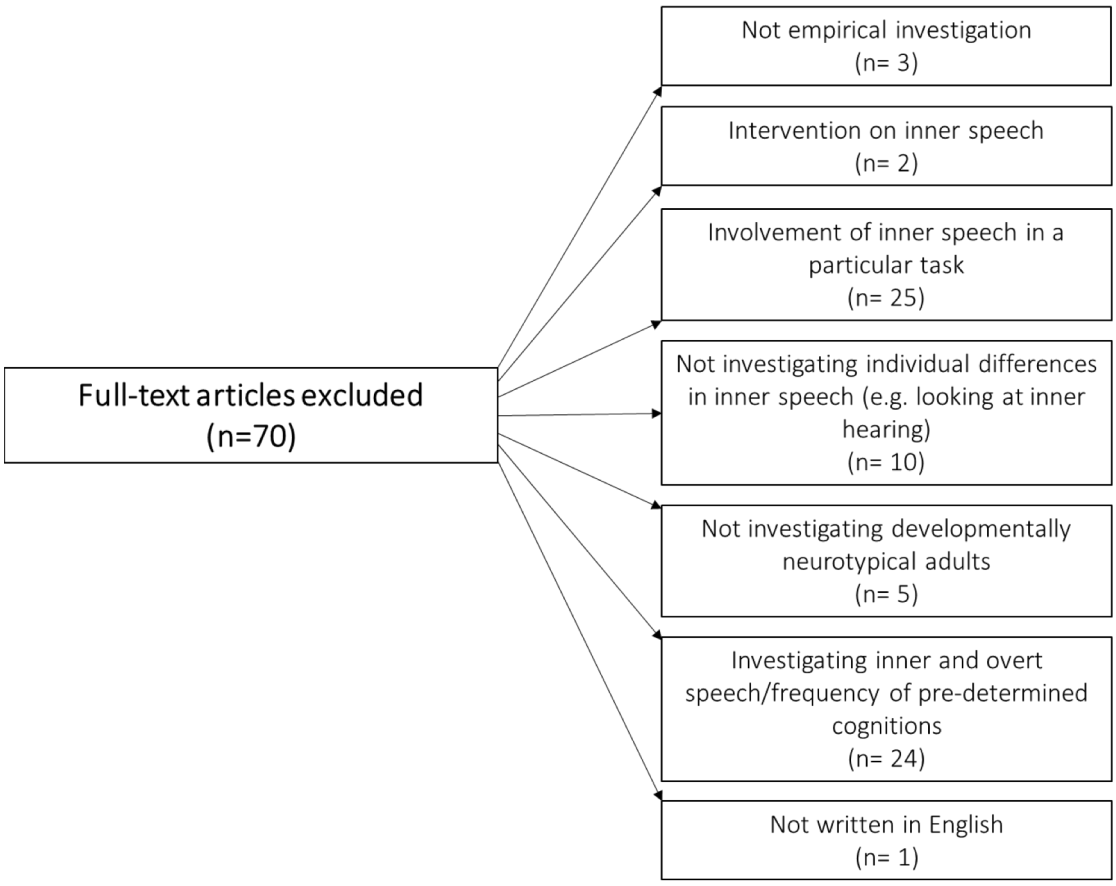


Figure 2_Reasons why full-text articles were excluded

A second rater also applied the inclusion/exclusion criteria to approximately 18 - 25% of the references at title, at abstract and at full paper (sees Figure 3 for numbers of references reviewed at each stage). References were randomly selected for double-rating (by using a random number generator in Excel and selecting papers in order from number 1 to the total that would be reviewed at that stage (i.e. up to 18 for paper)). Inter-rater reliability was assessed by Cohen's weighted kappa (as the categories of 'excluded', 'queried' and 'included' were considered ordered). This indicated that inter-rater reliability was 'good' for title (Weighted Kappa = 0.74), 'very good' for abstract (Weighted Kappa = 0.8) and 'good' for paper (Weighted Kappa = 0.77).

At all levels of screening (i.e. title, abstract and full-paper), there were discrepancies only between excluded and queried, or between queried and included. That is, for no reference did one rater think a paper should be included, and the other rater think the paper should be excluded. Discrepancies were resolved through discussion.

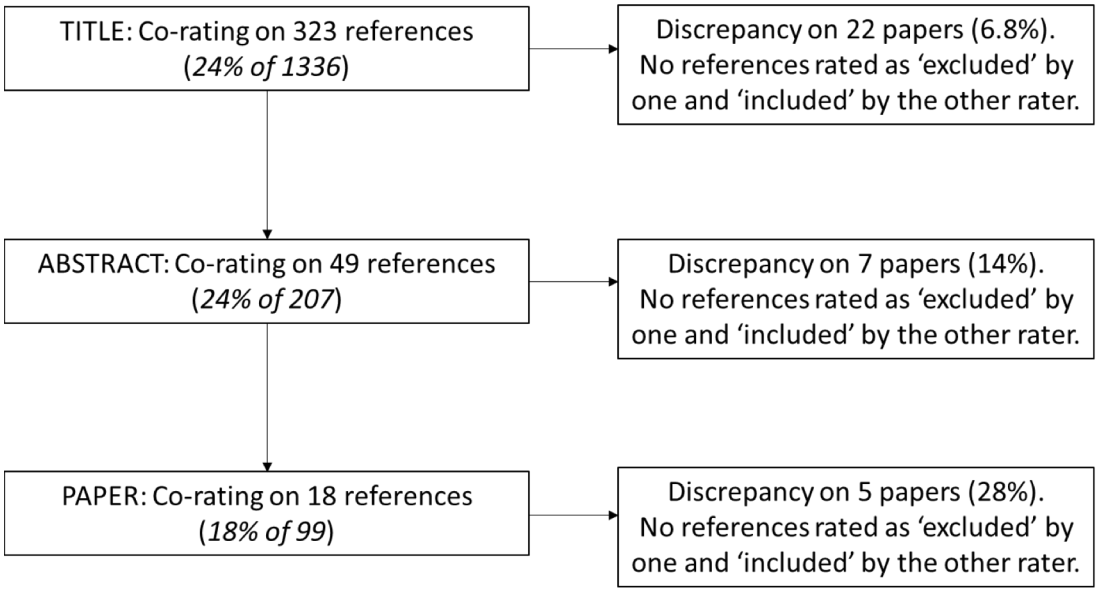


Figure 3_ Numbers of records reviewed by both raters at title, abstract and full-text.

Data Extraction

Inner speech measurement

The type of method used in each study was coded. Each method was coded on whether it was an objective or subjective measure and whether it relied on retrospective reporting of inner speech, as well as whether the inner speech measured was spontaneous or elicited by the researchers. Where possible, the authors' definition of inner speech was recorded, with particular attention paid to whether authors specified the concepts of 'verbal' and 'in head/without sound' (see quality assessment below). Inner speech features that were investigated by each method were coded and it was recorded whether each method was validated against another method.

Methods of measurement

The studies fell into five types of measurement methods. A sixth category included studies that had combined measurement methods. An overview of the study sample was reported for each of the papers, and different information was then extracted, depending on the type of method used:

Questionnaires and Interviews

The types of questions asked about inner speech were detailed. Brief notes on reliability and validity were reported, as were the method's findings.

Introspection

The protocol used in order to record participants' experiences was reported, as was the method the authors used to extract information from individuals' experiences and the categories of information that authors coded. If possible, inter-rater reliabilities specifically were reported, as well as brief notes on reliability and validity, and the method's findings.

Experimental

The protocol of the experimental method used was extracted, along with information on the stimuli used. Information about any controls used for confounding variables was reported if possible. The method's findings were reported.

Neuroimaging

The task that the study had used in order to elicit inner speech was reported, as were brain areas which were associated with elicited inner speech.

Lesion

The instructions and tasks used to examine inner speech functioning in individuals with lesions were presented, as well as findings from these experiments.

Combination methods

The methods used in these studies were reported as well as brief notes on reliability and validity, and the method's findings.

Quality Assessment Method

The quality of the included papers was assessed with respect to the method of measurement used to investigate inner speech, as well as the overall methodology of the paper.

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) is a checklist that can be used to rate the methodological quality of studies on measurement properties (Terwee et al., 2012). This checklist was used to rate the methodological quality of relevant measures in this review (i.e. studies that used questionnaire and interview methods) on the following measurement properties: internal consistency, reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity, responsiveness and interpretability (Mokkink et al., 2010). Measurement properties were rated as either 'Excellent', 'Good', 'Fair' or 'Poor' quality, in accordance with the COSMIN manual (Terwee et al., 2012).

As far as is known, no relevant, good quality assessment tools have been developed for systematic reviews of some methods of measurement included in this review (e.g. introspection) or for reviews of cross-sectional studies more generally (Sanderson, Tatt, & Higgins, 2007). Sanderson et al. (2007) conducted a systematic review of tools developed for assessing quality and susceptibility to bias in observational studies in epidemiology. They concluded that the most important domains for quality assessment include appropriate selection of participants, appropriate measurement of variables and appropriate control of confounding if investigating differences between groups. They also recommended that tools should provide a simple checklist, rather than a scale corresponding to varying levels of quality (Sanderson et al., 2007). Experts in the field of meta-analysis have also advised against using scale-based or aggregate ratings of study quality (Higgins et al., 2011).

The Agency for Healthcare Research and Quality (AHRQ) has investigated quality assessment of the evidence reported by observational studies, for the purpose of making evidence-based practice recommendations (Owens et al., 2010). Taylor, Hutton, and Dudley (2014) adapted AHRQ's recommendations for methodological quality criteria in order to develop an idiosyncratic quality assessment tool for their own systematic review. Their review aimed to investigate a perceptual bias (the 'jumping to conclusions' bias observed in some individuals with psychosis) which is assessed by using experimental tasks (Taylor et al., 2014). Taylor et al.'s (2014) tool included assessment of participant selection, variable measurement and control of confounding, the three most important domains for quality assessment as suggested by Sanderson et al. (2007). Because of the lack of a suitable quality assessment tool, relevant sections of Taylor et al.'s (2014) protocol were adapted for use with the current

systematic review, in order to assess the quality of each study's methodology. Taylor et al. (2014) suggested using the categories of 'Yes', 'No', 'Partially' and 'Can't tell' to grade the following questions:

- Adequate description of the cohort? - based on characteristics such as age, gender, ethnicity.
- Unbiased selection of the cohort? - based on inclusion/exclusion criteria and recruitment strategy.
- Selection minimises baseline differences in prognostic factors between groups (if appropriate)? - based on selection of the comparison group and matching of comparison to clinical group on key demographics.
- Missing data handling appropriately (if appropriate)? - based on whether missing data was reported, and if present, steps taken to minimise.
- Controlling for confounding (if appropriate)? - based on whether study identified and controlled for important confounding variables.

In addition, each study's definition of inner speech was rated for its quality. Definitions were rated according to their inclusion of two important concepts. These related to the verbal nature of inner speech, as well as the fact that it is silent and occurs in one's head. The following rating system was applied to the definitions:

- 'Excellent' - included well defined concepts of 'verbal' and 'in head'.
- 'Good' - included well defined concepts of either 'verbal' or 'in head'.
- 'Fair' - included different definitions or unclearly defined concepts of either 'verbal' or 'in head'.
- 'Poor' - if no definition was given.

Relationship between inner speech measurements and psychopathology

Each paper was then rated on whether it did or did not also investigate a psychological disorder or symptom. If the paper did investigate such phenomena, then information was extracted as to whether it was a clinical sample, as well as information about the psychological symptom(s) measured and whether the study found any relationships between

inner speech and the symptom(s). Papers that looked at differences in activation of brain areas between clinical groups and controls were not included in this section, as it was felt that differences in brain activity were not necessarily equivalent to individual differences in inner speech itself.

Results

Methods of assessing inner speech

In total 29 studies were included in this review (see Table 1).

Questionnaire and interview methods all relied on participants' subjective and retrospective reporting of their inner speech. Given that these methods relied on retrospective reporting, they were not considered to be assessing either elicited or spontaneous inner speech. These methods tended to provide information about the structure of inner speech, as well as frequency and content, to a lesser extent.

Introspective techniques all used subjective accounts, tended to rely on spontaneous inner speech and could be both retrospective and based on the present moment. These techniques coded the raw data collected through participant introspection on various categories to allow assessment of frequency, content and valence of inner speech.

Experimental methods tended to be both objective and subjective and were based on the present moment, but relied on elicited inner speech. They provided information about the structure of inner speech, or whether it was present or not.

If a definition of inner speech was provided for papers that used these three methods, it tended to refer to at least one of the key features of 'verbal' or 'in head'.

Neuroimaging techniques relied on non-retrospective and elicited inner speech. This method tended to produce objective information about the brain areas associated with elicited inner speech. However this method did not have an objective way of verifying that participants were actually engaging in inner speech during the brain scans, and relied on subjective self-report of inner experience. Definitions of inner speech were not typically reported.

Lesion studies were objective and relied on elicited, non-retrospective inner speech. They provided information about dissociations between inner speech and other cognitive abilities, as well as brain areas associated with this phenomenon.

Only two studies included in this review validated their methods of measurement against another method (seven percent). The methods used in these combination studies complemented each other on either the subjective and objective index (Kühn, Fernyhough, Alderson-Day, & Hurlburt, 2014), or on whether the measure was retrospective or relating to the present moment (Alderson-Day & Fernyhough, 2015). These papers' definitions of inner speech were of at least good quality.

Table 1. Characteristics of methods of measurement

Author, year	*Method*	*Name of method*	*Objective/ Subjective*	*Retrosp-ective?*	*Spontaneous or Elicited*	*Definition of Inner Speech*	*Quality of Definition*	*Characteristic measured*	*Validated against other method?*
McCarthy-Jones & Fernyhough, 2011	Q	Varieties of Inner Speech Questionnaire (VISQ)	Subjective	Yes	NA	Implication that IS is different to OS: "assumption that inner speech is comparable in form and structure to overt speech." p.1586.	Poor	Structure, Content	No
Alderson-Day et al., 2014	Q	VISQ	Subjective	Yes	NA	"the internal monologue"; "a silent form of speech" p.288	Excellent	Structure, Content	No
Delamillieure et al., 2010	Q/Iv	The Resting State Questionnaire (RSQ)	Subjective	Yes	NA	"Inner speech corresponds to imagined speech (talking to oneself with one's voice) without overt production." p566	Excellent	Frequency, Content (past/ future).	No
Langdon et al., 2009	Iv	[Author developed interview]	Subjective	Yes	NA	"inner-speech as 'thinking in words' … 'verbal thought'" … and thinking … is the act of using language to talk to oneself internally" p. 655	Good	Structure, Content, Frequency	No
Hurlburt & Heavey, 2002	Ins/Iv	Descriptive Experience Sampling (DES)	Subjective	No	Spontaneous	Not reported	Poor	Frequency	No
Hurlburt et al., Heavey, 2002	Ins/Iv	DES	Subjective	No	Spontaneous	"the experience of speaking words in the person's own voice, with the same vocal characteristics … as the person's own external speech, but with no external (real) noise" p. 128	Excellent	Frequency	No

Table 1. Characteristics of methods of measurement

Author, year	*Method*	*Name of method*	*Objective/ Subjective*	*Retrosp-ective?*	*Spontaneous or Elicited*	*Definition of Inner Speech*	*Quality of Definition*	*Characteristic measured*	*Validated against other method?*
Heavey & Hurlburt, 2008	Ins/Iv	DES	Subjective	No	Spontaneous	"speaking words in your own voice, usually with the same vocal characteristics as your own external speech, but with no external (real) sound or motion" p. 802	Excellent	Frequency	No
Morin et al., 2011	Ins	[Thought listing]	Subjective	Yes	NA	“normal inner speech” p. 1714	Poor	Content	No
Payne & Manning, 1991	Ins	[Random sampling method]	Subjective	Yes	NA	"self-talk or internal dialogue" p. 48	Fair	Content, Valence	No
Lyxell et al., 1994	E	[Word Judgements]	Objective	No	Elicited	"inner speech refers to activities where some form of existing internal mental representation of sound is used for various purposes" p 179.	Good	Structure	No
Oppenheim & Dell, 2008	E	[Word judgements]	Subjective	No	Elicited	"we produce inner speech the same way that we speak, except that articulation is not present … we hear the speech in our mind" p. 529	Excellent	Structure	No
Korba, 1990	E	EMG	Objective	No	Elicited	"the private construction of elliptical thought (inner speech) occurs in a linguistic form independent of the semantic, syntactic and pragmatic rules of external speech" p. 1043	Good	Presence or absence	No
Livesay et al., 1996	E	EMG	Objective	No	Elicited	"mentally engage in … language recitation" p. 1357	Good	Presence or absence	No

Table 1. Characteristics of methods of measurement

Author, year	*Method*	*Name of method*	*Objective/ Subjective*	*Retrosp-ective?*	*Spontaneous or Elicited*	*Definition of Inner Speech*	*Quality of Definition*	*Characteristic measured*	*Validated against other method?*
Hinke et al., 1993	N	fMRI	Both	No	Elicited	"speaking in one's mind without physical articulation or vocalisation" p. 677-678	Excellent	Brain areas associated with IS	No
Baciu et al., 1999	N	fMRI	Both	No	Elicited	None reported	Poor	Brain areas associated with IS	No
Lee et al., 2009	N	fMRI	Both	No	Elicited	None reported	Poor	Presence or absence of elicited IS?	No
Abe et al., 2011	N	fMRI (ICA)	Both	No	Elicited	None reported	Poor	Brain areas associated with IS	No
Scholkmann et al., 2013	N	Functional NIRS	Both	No	Elicited	None reported	Poor	Effect of IS on cerebral hemodynamics	No
Scholkmann et al., 2014	N	Functional NIRS and Capnography	Both	No	Elicited	"inner speech, i.e., speech not spoken aloud" p 017002-1	Excellent	Effect of IS on cerebral hemodynamics	No
Shergill et al., 2003	N	fMRI	Both	No	Elicited	None reported	Poor	Brain areas associated with IS	No
Simons et al., 2010	N	fMRI	Both	No	Elicited	"inner speech or the silent articulation of words" p232	Good	Brain areas associated with IS	No

Table 1. Characteristics of methods of measurement

Author, year	*Method*	*Name of method*	*Objective/ Subjective*	*Retrosp-ective?*	*Spontaneous or Elicited*	*Definition of Inner Speech*	*Quality of Definition*	*Characteristic measured*	*Validated against other method?*
Vercammen et al., 2011	N	fMRI	Both	No	Elicited	"Inner speech, sometimes referred to as "verbal thoughts" or "verbal imagery" has been defined as our ability to "talk" silently to ourselves" p 1009	Excellent	Brain areas associated with IS	No
Rapin et al., 2012	N	fMRI	Both	No	Elicited	"private, self-generated thoughts"	Fair	Brain areas associated with IS	No
Ford et al., 2014	N	ERP	Both	No	Elicited	None reported	Poor	Brain areas associated with IS	No
Levine et al., 1982	L	[Neuropsychological tasks]	Objective	No	NA	"First, it refers to the subjective phenomenon of talking to oneself … without uttering a sound. Second it refers to the objectively measurable ability to appreciate the auditory-articulatory structure of speech irrespective of its meaning." p391.	Excellent	Dissociation between IS and other cognitive abilities.	No
Geva, Jones, et al., 2011	L	voxel-based lesion-symptom mapping	Objective	No	Elicited	"the ability to speak silently in one's head" p 3071.	Excellent	Brain areas associated with IS	No
Geva, Bennett, et al., 2011	L	[Verbal judgement tasks]	Objective	No	Elicited	"Inner speech is the ability to talk to yourself in your head and listen to what you are saying" p. 323	Excellent	Dissociation between IS and other cognitive abilities.	No

Table 1. Characteristics of methods of measurement

Author, year	*Method*	*Name of method*	*Objective/ Subjective*	*Retrosp-ective?*	*Spontaneous or Elicited*	*Definition of Inner Speech*	*Quality of Definition*	*Characteristic measured*	*Validated against other method?*
Kühn et al., 2014	C (Ins, N)	DES, fMRI	Objective (N) and Subjective (Ins)	No	Spontaneous	"aspects of inner experience" p.1; "inner speaking" p. 2	Good	Brain areas associated with IS, Frequency	Yes
Alderson-Day & Fernyhough, 2015	C (Q, Ins	VISQ, Random sampling method	Subjective	Yes (Q) and no (Ins)	NA	"Inner speech – talking to oneself silently and internally" p.1	Excellent	Structure, Content, Valence	Yes

Notes: IS – Inner Speech. Q – Questionnaire. Iv – Interview. Ins – Introspection. E - Experiment. N – neuroimaging. L – Lesion. C – Combination, EMG – electromyographic. fMRI – functional Magnetic Resonance Imaging. ICA – independent component analysis. NIRS – near-infrared spectroscopy. ERP – event related potential.

Assessment methods: methodological quality of paper and findings

The six categories of methods, their findings and observations about the particular paper's quality will be presented separately. In the table headings (Tables 2, 4, 5, 6, 7 and 8), 'Description' refers to whether the sample was adequately described, 'Selection' refers to whether there was unbiased selection of the participants and 'Min Difs' refers to whether authors attempted to minimise baseline differences between the groups (see Appendix 4 for information on the study sample and selection procedures that were used to make these decisions).

None of the papers reviewed provided prospective sample size calculations.

Questionnaires and interviews

Table 2 presents papers that used questionnaire and interview methodologies. Papers tended to adequately describe the sample with regard to age, gender and disease characteristics (if relevant), although none reported participant ethnicity (see Appendix 4 for more information). Three of the four papers did not adequately report on both the selection procedure and inclusion/exclusion criteria, so it was not possible to determine whether there had been unbiased selection of the sample. If the measure had been used in two samples, these tended to be of a similar demographic.

McCarthy-Jones and Fernyhough (2011) designed the Varieties of Inner Speech Questionnaire (VISQ) to measure four aspects of inner speech. These include 'Dialogic' which refers to the interplay between two perspectives (e.g. "I talk back and forward to myself in my mind about things"), 'Condensed' which examines the extent to which inner speech is abbreviated (e.g. "I think to myself in words using full sentences"), 'Other people' which asks about the presence of others' voices in inner speech (e.g. "I experience the voices of other people asking me questions in my head"), and 'Evaluative/Motivational' which investigates the role of inner speech in evaluating situations, people and the self (e.g. "I think to myself in inner speech about what I have done, and whether it was right or not"). Delamillieure et al. (2010) designed the Resting State Questionnaire (RSQ) (which also included an interview), which was administered to individuals following a resting-state functional magnetic resonance imaging

(fMRI) scan. Participants rated the amount of time they spent engaged in different forms of inner experience, which included 'inner language', a combination of inner speech and auditory mental imagery. Participants were asked whether the content of each mental activity related to ongoing learning activities, memory reminiscences, or prospective thoughts, and whether it was associated with a positive or negative valence. They also answered questions about the content of these experiences, such as whether inner speech was related to objects, places or people. Langdon et al. (2009) developed an interview to ask about participants' inner speech experiences, which included questions about the frequency, structure, speed and content of their inner speech.

Specific findings were that evaluative/motivational and dialogic inner speech were endorsed more frequently than condensed inner speech, or experiencing the presence of others' voices (Alderson-Day et al., 2014; McCarthy-Jones & Fernyhough, 2011). Individuals reported that over half of their inner speech occurrences were related to either positive or negative emotions, and individuals who had received more years of education tended to engage in more prospective inner speech (Delamillieure et al., 2010). Individual variation was found by all measures used, indicating that these methods allow for empirical investigations of inner speech differences.

Table 2. Quality, content and findings of questionnaire and interview measures

Author, year	Sample Overview	Description	Selection	Min Difs	Questions	Reliability/validity (see Table 3)	Missing data	Findings
McCarthy-Jones & Fernyhough, 2011	Two samples of university students	Yes (except ethnicity)	Can't tell	NA	Developed through discussions with students and informal piloting. 18 items, 4 subscales: 'Dialogic', 'Condensed', 'Other people', 'Evaluative/Motivational'. Scored on Likert scale (1: 'certainly does not apply to me' through 6: 'certainly does apply to me').	Had reliable 4 factor structure.	Yes	82% experienced some evaluative/motivational, 77% experienced dialogic, 36% endorsed condensed and 26% endorsed others in speech.
Alderson-Day et al., 2014	University students	Yes (except ethnicity)	Can't tell	NA	As above – used 18 item VISQ.	Found 4 subscales had good reliability.	Can't tell	Dialogic scores ranged between 4 and 24 (M=18), Evaluative between 5 and 24 (M=19), Other people between 5 and 30 (M=12), Condensed between 5 and 27 (M=14).
Delamillieure et al., 2010	Healthy volunteers	Yes (except ethnicity)	Can't tell	NA	Post resting state scan, rate time spent on different types of inner experience – inner language was one phenomenon. Information elicited regarding temporal focus and valence of experience as well as content. At end, rated time spent on each mental activity again.	No reliability statistics reported, only conducted in one sample. Test-retest on time spent on mental activities (not whole measure) 10 minutes later.	Can't tell	30% engaged in inner language (including both IS and auditory hearing). Positive or negative emotions reported for 62% of IS occurrences. Significant effect of education level with prospective IS thoughts more frequent in people with higher education level.
Langdon et al., 2009	Patients with SZ and controls	Yes (except ethnicity)	Yes	Yes	Interview asked about frequency of IS, structure of IS, speed and questions relating to content.	Interview was "completely new". No reliability statistics reported.	Can't tell	Results related to differences between people with AVHs and controls and so will be reported below.

Notes: SZ – schizophrenia, AVH – auditory verbal hallucinations

Quality assessment of questionnaires using COSMIN

Table 3 presents COSMIN ratings of measurement properties of the questionnaire and interviews above (see Appendix 5 for information used to establish a quality rating). The VISQ (McCarthy-Jones & Fernyhough, 2011), used by McCarthy-Jones and Fernyhough (2011), Alderson-Day et al (2014) and Alderson-Day and Fernyhough (2015), was found to have very good measurement properties according to the COSMIN checklist (Terwee et al., 2012). The RSQ and interview (Delamillieure et al., 2010), had much poorer properties as evaluated by this checklist, and it was not possible to assess many of the psychometric properties of the interview developed by Langdon et al. (2009).

Table 3. COSMIN ratings of questionnaires and interview measures

Author, year	Missing data acceptable? (Less than 20%)	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypotheses testing	Cross-cultural validity	Criterion validity	Responsiveness	Interpretability
McCarthy-Jones & Fernyhough, 2011	Yes	Excellent	Excellent	Excellent	Good	Excellent	Fair	NA	NA	NA	Excellent
Delamillieure et al., 2010	Not reported	Poor	Poor	Poor	Fair	NA	Poor	NA	NA	NA	Fair
Langdon et al., 2009	Not reported	Poor	Poor	Poor	Fair	NA	Good	NA	NA	NA	Good

Introspection

Table 4 presents information about the introspective methods. Studies were varied with regard to adequate description of the sample (see Appendix 4), with one study failing to report participant's ages (Payne & Manning, 1991). None of the studies reported ethnicity. The papers also tended not to report their recruitment strategies or inclusion and exclusion criteria, making it difficult to assess for the presence of bias in sample selection. However when information was available bias was not found (Hurlburt & Heavey, 2002).

In Descriptive Experience Sampling (DES) (Hurlburt & Heavey, 2002) participants carry a beeper in their natural environments and make notes about their inner experience when prompted by a beep. Up to 16 characteristics of inner experience were identified in the studies included in this review, of which 'inner speech' was consistently identified as one of the five most common experiences (along with 'image' and 'feeling' for example) (Heavey & Hurlburt, 2008; Hurlburt & Heavey, 2002; Hurlburt, Koch, & Heavey, 2002).

Morin et al. (2011, p. 1715) asked participants to list the self-verbalisations they had had over the previous week. They developed a coding scheme for the raw data, which included categories such as whether inner speech focused on the self, the environment or activities (see Table 4). It was not clear why these categories were chosen. Payne and Manning (1991, p. 49) asked participants to keep logs of their self-talk (described as "what they said to themselves about situations") for two days. This raw data was coded on three dimensions relating to the transactional state of the self-talk (that is, whether it included an 'adult', 'parent' or 'child' state (Berne, 1968)), whether it had an internal or external locus of control and whether it had a positive, negative or neutral valence.

The introspective method produces raw data, which can then be examined by investigators for the presence of particular categories of information. Inter-rater reliabilities with regard to these coding categories tended to be good, when this information was reported.

Individual differences were observed using these methods. For example Hurlburt et al. found that individuals whose external speech was very fast, tended to have less 'pure' inner speech (without other cognitive components such as images) than those who spoke at a typical rate, and Heavey and Hurlburt (2008) found that only a quarter of their sample engaged in inner speech.

Table 4. Quality, content and findings of introspective methods

Author, year	Sample Overview	Description	Selection	Min Difs	Protocol	Coding categories	Inter-rater on coding?	Missing data	Reliability/validity	Findings
Hurlburt & Heavey, 2002	University students	Yes (except ethnicity)	Not clear	NA	Beeper in environment, over several days. 6 samples from each participant.	Characteristics (including inner speech) extracted. Interviewers divided between, and all samples coded by, 2 independent authors.	Yes. Interobserver agreement on 5 most frequent = 91%.		Both interviewers coded 5 characteristics (including inner speech) in at least 15 of 60 samples. Inner speech interobserver reliability (Kappa) = .88.	Demonstrates inter-rater reliability.
Hurlburt et al., 2002	University students. High-speech rate and controls	Yes (except ethnicity)	Yes	Yes	Beeper in environment. Sampling over 3-5 days.	Investigators extracted characteristics (including inner speech)'.	Not reported	Not reported	References Hurlburt and Heavey (2002).	5 times more simple speech (IS without other inner experiences – e.g. image) in control group. Conclude high speed have less inner speech.
Heavey & Hurlburt, 2008	University students	Yes	Not clear	NA	Beeper in environment. 6 beeps/day, for 3 days.	Inner speech, 1 of 16 phenomena coded.	No	Yes, no subject > 20%	Method not yet replicated by other labs and triangulated with other methods of inquiry.	Inner speech present in approximately one quarter of sampled moments
Morin et al., 2011	University students	Yes (except ethnicity)	Not clear	NA	List self-directed thoughts from week.	Coding scheme categories A: self or others, B: functions, C: people, D: activities; E: environment; F: events.	Second coder on subset of 100 reports. Inter-rater reliability not reported.	N/A	"exhibits good criterion-related, concurrent and discriminant validity", p. 1715 (without reference). Non-inner speech occurrences may have been listed.	Participants mostly report talking to themselves about themselves, as well as about future events.
Payne & Manning, 1991	Pre-service teachers	No (no age)	Not clear	NA	Kept self-talk logs for two 24 hr periods.	Rated on transactional state locus of control, directional state.	2 independent raters obtained 92% inter-rater reliability.	N/A	Face validity, no other information reported.	Engaged in a majority of negative (50%), child oriented (67%), external (75%) self-talk.

Experimental tasks

Table 5 presents information about the papers that used experimental tasks. Half of the studies failed to adequately describe the study sample in terms of age (Lyxell et al., 1994) and gender (Oppenheim & Dell, 2008), and all failed to report participant ethnicity (see Appendix 4 for more information). It was neither possible to determine whether the sample selections were unbiased due to a lack of information in the papers, nor to determine whether baseline differences between groups had been minimised. When applicable, control of confounding variables was found to be suitable (Oppenheim & Dell, 2008).

Lyxell et al. (1994) asked participants to silently decide whether two words rhymed, while Oppenheim and Dell (2008) gave participant's tongue-twisters to recite in inner speech, and asked them to report any errors. Tongue twisters can elicit a lexical bias effect (tendency for errors to result in words over non-words) and a phonemic similarity effect (where similar phonemes are more likely to interact in slips than less similar phonemes) (Oppenheim & Dell, 2008).

Findings from these silent word judgement tasks relate to the structure of inner speech. For example Oppenheim and Dell (2008) found that the phonemic similarity effect was only present in external slips, while the lexical bias was present during both inner and external speech decisions. They concluded that inner speech was impoverished in comparison to external speech. It should be noted that this method relied on subjective reporting of inner speech errors.

Korba (1990) and Livesay et al. (1996) used electromyography (EMG) to measure participants' lip movements when they engaged in elicited silent speech. These studies indicate that it is possible to determine whether an individual is engaging in elicited, but not spontaneous, inner speech (Korba, 1990; Livesay et al., 1996). Therefore it appears that at least some form of inner speech is observable when using this technique.

Table 5. Quality, content and findings of experimental methods

Author, year	Sample Overview	Description	Selection	Min Difs	Experimental Protocol	Stimuli	Control for confounding variables?	Findings
Lyxell et al., 1994	Deafened adults controls	No (age, ethnicity)	Not clear	Not clear	Computer task – rhyme judgement. No practice trial was described.	Word pairs rhymed, did not rhyme and were orthographically similar, or did not rhyme and were orthographically dissimilar.	Yes – no significant difs between groups on speed of reading	Deafened adults performed less accurately on non-rhyming orthographically similar words pairs (i.e. similar spelling). Decreased accuracy associated with more years since became deaf. Deafened subjects appear to less capable of forming internal phonological representations of words.
Oppenheim & Dell, 2008	Local residents	No (gender, ethnicity)	Not clear	Not clear	Tongue-twisters used to produce 'slips'. Recited sequence aloud 4 times, then cued to recite aloud or internally 4 times and subjectively report any errors.	4 word stimuli presented on computer (e.g. 'bean reed reef beach').	N/A	Lexical bias was present in both inner and overt slips, phonemic similarity effect only present in external slips. Concluded that inner speech is impoverished compared to external speech.
Korba, 1990	University students	Yes (except ethnicity)	Not clear	NA	EMG recordings of lip movements, while attempting to solve verbal task in inner speech. Once solved, recorded inner speech during task (elliptical word count), and then explained in full (extended word count).	Not reported	N/A	Neither the elliptical word count, nor extended word count correlated with duration of EMG recorded lip movements. Faster problem solving led to strongest correlation for eliptical (i.e. elliptical - r=.29, p = .07; extended r=.09, p= .58). 4,000 words/minute as estimate of IS speed.
Livesay et al., 1996	Volunteers from a university setting	Yes (except ethnicity)	Not clear	NA	EMG recordings of lip and forearm movements during inner speech and visualisation tasks.	IS - silent recitation of Pledge of Allegiance. Visualisation - 'visualise in your mind's eye the American flag'.	N/A	During IS task mean lip EMG activity increased significantly from rest (t=-2.88, p>0.01). Lip activity was not present during non-language visualisation condition.

Imaging

Table 6 presents information about the tasks used to investigate inner speech in neuroimaging studies. All papers adequately described the sample in terms of age and gender, although none reported ethnicity (see Appendix 4 for more information). The majority of authors did not give information on sample recruitment or selection criteria, so it was not possible to assess whether selection of the sample was biased. If information was reported, selection appeared to be partially biased. Papers tended to adequately minimise baseline differences between groups of individuals with a diagnosis of schizophrenia and controls.

All tasks were based on asking participants to say something silently, and so relied on elicited inner speech. Papers did note that it was not possible to determine exactly what inner experience individuals were engaging in during these tasks (e.g. people may also have images (Lee et al., 2009)). Simons et al. (2010) noted that 'intrinsic inner speech' may occur during baseline tasks, indicating that they considered this to be a different phenomenon from the elicited inner speech they were investigating.

Brain areas that were found to be associated with the production of elicited inner speech included the inferior frontal gyrus (which includes Broca's area), particularly in the left hemisphere, as well as the anterior cingulate gyrus, the insula and the medial temporal gyrus (see Table 6).

Table 6. Quality, content and findings of neuroimaging methods

Author, year	Sample Overview	Description	Selection	Min Difs	Task	Brain Areas/Findings
Hinke et al., 1993	Healthy participants	Yes (except ethnicity)	Not clear	NA	Silent word generation from the alphabet (i.e. A-apple, B-boat, C-car etc).	More activation in L Broca's area compared with its right-side homologue. Expected as all participant's right handed and so expect left hemisphere language dominance.
Baciu et al., 1999	Healthy participants	Yes (except ethnicity)	Not clear	NA	Asked to imagine scenes (e.g. office, street) and generate nouns of objects in scenes, covert and overt conditions.	Same areas (particularly Broca's area) activated during both conditions (except for primary motor area – for overt only). Left hemispheric dominance (of right handed participants) found for both covert and overt.
Lee et al., 2009	Healthy participants	Yes (except ethnicity)	Not clear	NA	6 tasks completed: right/left hand and right foot motor imagery, mental calculation, visual imagery and inner speech. IS task: silently speak sentences of choice (although use same sentence once chosen). Repeated tasks during a second scan.	Produced task-specific regions of interest (ROIs) for each participant. Areas specifically activated by IS not reported. Found variation in spatial distribution and size of ROI between subjects. Mean classification performance across subjects was 78% (SD = 16%) for scan 1 and 71% (SD = 20%) for scan 2.
Abe et al., 2011	Healthy participants	Yes (except ethnicity)	Not clear	NA	Participants selected tongue twister for responding. Tasks followed this pattern - select number from 1 - 10 (wrote on paper at beginning and end of scan) - in scan played numbers 1 - 10 twice in pseudorandom order and when heard number asked to repeat tongue twister silently in mind till heard next number.	Focus on Broca's area (left frontal operculum), but for 6 of 8 participants, signals also extended caudally to Rolandic operculum (implicated in syntactic encoding). Suggests that occasional errors in syntax or attempts not to make errors, activate networks for syntactic processing even in silent speech
Scholkmann et al., 2013	Healthy participants	Yes (except ethnicity)	Not clear	NA	Speech therapist recited respective either hexameter or prose text and participant repeated with inner speech. Also completed mental arithmetic. Each task performed on separate day.	Inner speech effects cerebral hemodynamics and oxygenation primarily by changes in arterial $PaCO_2$ caused by variations in respiration and secondarily by increased neuronal activity of the pre-frontal cortex.
Scholkmann et al., 2014	Healthy participants	Yes (except ethnicity)	Not clear	NA	6 Tasks: 2 listening to live recitation, 2 listening to recorded recitation, 2 inner speech (of a hexameter and of alliterative text).	A priori focus on left and right pre-frontal cortex. Inner speech caused changes in $PaCO_2$.

Table 6. Quality, content and findings of neuroimaging methods

Author, year	Sample Overview	Description	Selection	Min Difs	Task	Brain Areas/Findings
Shergill et al., 2003	Patients with SZ and controls	Yes (except ethnicity)	Not clear	Yes	Participants covertly generated the word 'rest' repeatedly at 2 self-paced rates (one every 1s, one ever 4s).	Increasing the rate of covert articulation was associated with activation in the left inferior frontal cortex.
Simons et al., 2010	Patients with SZ and controls	Yes (except ethnicity)	Not clear	Yes	Listened to sentence, then either listened to sentence again (listening task) or repeated same sentence covertly (inner speech task).	Increased activation of left inferior frontal gyrus and anterior cingulate gyrus during inner speech compared to listening condition across groups.
Vercammen et al., 2011	Patients with SZ	Yes (except ethnicity)	Partially	NA	Shown bisyllabic Dutch words. "Semantic condition" - decide whether word positive or negative, "phonology condition" – judge metrical stress (is stress on first or second syllable), and baseline condition (fixation cross).	ROIs active during phonology in contrast to baseline included Bilateral inferior frontal gyrus, opercular and traingular parts and bilateral superior and medial temporal gyrus, insula, supplementary motor area, angular and supramarginal gyrus and anterior cingulate cortex.
Rapin et al., 2012	Patients with SZ and controls	Yes (except ethnicity)	Partially	Yes	IS task: presented with picture of an object and asked to silently generate definition of this object; Speech perception task: presented with picture of an object, while listening definition of the object.	Functional network involved primarily in silent thought generation displayed strong dorsal anterior cingulate cortex activation.
Ford et al., 2014	Patients with SZ and controls	Yes (except ethnicity)	Not clear	Partially	IS task - repeat statements silently, listening task – listen to recordings of statements in own voice and baseline - in silence (when spontaneous inner speech could occur).	Focus only on auditory cortex – which was activated during directed inner speech task, but was reduced in comparison to listening condition.

Notes: SZ – schizophrenia, $PaCO_2$- arterial CO_2 pressure. $P_{ET}CO_2$ - end-tidal CO_2, a reliable and accurate estimate of $PaCO_2$.

Lesion

Table 7 presents the lesion studies included in this review. All papers adequately described the sample in terms of gender and age, although ethnicity was not reported (see Appendix 4). Information was lacking on recruitment procedures so it was not possible to assess whether the papers selected their samples without bias. The one paper which investigated group differences was given a rating of partially minimising baseline differences between groups, as although age was matched, it was not certain whether gender was matched or reading ability assessed (Geva, Bennett, et al., 2011).

Most of the methods used in the lesion studies were similar to the experimental paradigms used in the above literature. However these lesion studies allow investigation into the relationships between inner speech and other cognitive functions. It appears that language ability does not depend on inner speech ability (Levine et al., 1982). In addition some participants with aphasia had preserved inner speech, but impaired overt speech, indicating that inner speech is not simply overt speech without articulation (Geva, Bennett, et al., 2011). Brain areas important for engagement in elicited inner speech included the left inferior frontal gyrus (Geva, Jones, et al., 2011).

Table 7. Quality, content and findings of lesion studies

Author, year	Sample Overview	Description	Selection	Min Difs	Instructions/Tasks	Findings
Levine et al., 1982	Individual without subjective inner speech following stroke	Yes (except ethnicity)	NA	NA	Neuropsychological tests	Participant had complete absence of speech production. His performance on other language tasks (reading, writing) were abnormal due to slowness, but accuracy was far above average.
Geva et al., 2011	Patients with aphasia	Yes (except ethnicity)	Partially	NA	Inner speech tasks (adapted from aphasia assessment (Kay, Coltheart, & Lesser, 1992)), during fMRI scan. Rhyme determination task: word pairs either orthographically similar (town-gown v hush-bush) or dissimilar (chair-bear v bond-hand). Homophone determination task (might-mite - yes, ear-oar - no). Practice trials to ensure no sound production.	Lesions to the inferior frontal gyrus pars opercularis (BA 44) and supramarginal gyrus (BA 40) and its adjacent white matter correlated with performance on rhyming task and similar (weaker) effects with homophone judgements.
Geva et al., 2011	Patients with aphasia and controls	Yes (except ethnicity)	Partially	Partially	Tasks performed during inner and overt speech conditions: Rhyme determination task, homophone judgement on real words and non-words (e.g. "zole" and "zole", v "noal" and "nool").	Patients as a group impaired relative to controls in all three tasks. Patient group is heterogeneous with IS abilities covering entire range from normal performance to severe impairment. Some patients have preserved IS and impairment in overt speech (OS), some had preserved OS and impaired IS.

Combination methods

Papers that used a combination of methods to investigate inner speech are presented in Table 8. Both studies described the samples in terms of age and gender but not ethnicity (Alderson-Day & Fernyhough, 2015; Kühn et al., 2014). It was not possible to determine whether sample selection was biased due to a lack of information about recruitment strategies (Kühn et al., 2014) and inclusion criteria (Alderson-Day & Fernyhough, 2015; Kühn et al., 2014).

Kühn et al. (2014) used a combination of fMRI and DES to investigate inner speech in one participant. This participant engaged in an fMRI task that elicited inner speech. She was trained on using DES in her natural environment, and then completed DES during a second fMRI scan, with an interview after the scan was completed. The experimenters coded the raw introspective data from DES samples in the scanner, for the presence of different inner experiences (including inner speech), and classified each sample into 'verbal', 'visual', 'bodily' and 'auditory'. Elicited inner speech was associated with activity in the same brain areas as in previous studies (see Table 6; e.g. Abe et al., 2011), including the left inferior frontal gyrus, superior temporal sulcus, and superior and middle temporal gyri. During the DES scan, samples that were coded as including inner speech were also associated with activity in the left inferior frontal gyrus, when compared to samples coded as other inner experiences.

This finding could be "understood as a validation of DES … [while] the DES descriptions lend support to fMRI techniques as a way of investigating" inner experience (Kühn et al., 2014, p. 6). An additional finding was that, compared with unprompted inner speaking as classified by DES, elicited inner speech was associated with activation in a larger network of brain areas. This study had only one participant, so caution must be taken with generalising its results.

Alderson-Day and Fernyhough (2015) administered the full VISQ to an undergraduate sample, before and after participants completed experience sampling via a smartphone app. Participants were randomly prompted by the app during their day and were asked to complete one question from each of the four VISQ subscales, and record whether they were thinking about the past, present or future. There were correlations between random sampling and questionnaire self-report with regard to the Condensed and Other People VISQ subscales, but not between the Dialogic and Evaluative subscales. In addition, subjective endorsement was lower during experience sampling than during questionnaire self-report.

Table 8. Quality, content and findings of combination methods

Author, year	Sample Overview	Description	Selection	Min Difs	Methods	Reliability/validity	Missing data	Findings
Kühn et al., 2014	Not clear	Yes (except ethnicity)	No	NA	Phase 1: fMRI elicited tasks – imagine saying words (e.g. 'pencil') v fixation cross. Phase 2: natural environment DES (see Table 4) – 6 random beeps a day for 4 days. Phase 3: in scanner DES. 4 beeps during 25 min scan for 9 scans. After beep wrote notes on pad (viewed by mirror on lap) and had interview after scan. DES coded for presence of 5 most common phenomena, including IS, also classified into 'verbal', 'visual', 'bodily' and 'auditory'.	Combination of phenomenology and neurophysiology – understood as both validation of DES, and to fMRI techniques. Wider spread of activation associated with imagined inner speech (Phase 1), compared to unprompted inner speaking (Phase 3).	No	Phase 1: elicited task v baseline – activated left IFG, STS and superior/middle temporal gyrus. Phase 3: brain activity classified as verbal showed activation of IFG. Brain activity modelled across all DES samples. 8 DES samples classified as 'inner speaking' occurring – models averaged and compared to average of remaining 28 DES samples – indicated activity in left IFG
Alderson-Day & Fernyhough, 2015	University students	Yes (except ethnicity)	Can't tell	NA	Used 18 item VISQ (see Table 2). Administered measure before and after sampling. Used experience sampling via smartphone app (*Inner Life*) – 2 random prompts a day. Inner speech questions = 1 item for each of 4 subscales (highest loading from original VISQ analysis – see Table 2). Also recorded whether thinking about past/present/future.	Significant correlations between random sampling and VISQ Condensed ($r = 0.69$, $p < 0.001$) and Other people ($r = 0.46$, $p < 0.001$), but not Dilaogic (after using adjusted alpha – $r = 0.3$, $p = 0.03$) and Evaluative ($r = 0.03$, $p = 0.85$). Test-retest reliability on VISQ acceptable for Dialogic ($r = 0.81$, $p < 0.001$), Condensed ($r = 0.88$, $p < 0.001$), Other People ($r = 0.68$, $p < 0.001$) and modest for Evaluative ($r = 0.42$, $p = 0.01$)	Yes VISQ 2 completed by 71%.	Endorsement of subscales were significantly lower when sampling method used, v questionnaire. Temporal thinking (high score - future, low score – past) associated with VISQ 1 Evaluative ($r = -0.32$, $p = 0.02$) and Condensed ($r = 0.36$, $p = 0.01$) also sampling Condensed ($r = 0.41$, $p = 0.003$).

Notes: fMRI – functional magnetic resonance imaging, IFG – inferior frontal gyrus, STS – superior temporal sulcus

Methods of investigating inner speech and psychopathology

Three of the studies included in this systematic review related variation in inner speech to variation in psychological symptoms of mental ill-health in non-clinical samples (see Table 9). Increased engagement in evaluative and motivational inner speech was associated with increased levels of anxiety as measured by questionnaire self-report (Alderson-Day & Fernyhough, 2015; McCarthy-Jones & Fernyhough, 2011), but not when measured by in the moment sampling (Alderson-Day & Fernyhough, 2015). Self-reported anxiety was also associated with the presence of other people in one's inner speech as assessed by both questionnaires and experience sampling (Alderson-Day & Fernyhough, 2015; McCarthy-Jones & Fernyhough, 2011).

More evaluative inner speech, and the presence of other's voices were associated with lower self-esteem (Alderson-Day et al., 2014). Dialogic, evaluative and others' voices in inner speech were related to an individual's proneness to experience auditory hallucinations (Alderson-Day et al., 2014; McCarthy-Jones & Fernyhough, 2011). Experiencing more adult perspectives (e.g. "how can I make the best of this situation"), and fewer child perspectives (e.g. "I'm so mad, I'm not going to teach this class") in one's inner speech was related to a greater internal locus of control and higher self-esteem (Payne & Manning, 1991).

Table 9. Inner speech differences and psychopathology

Author, year	Clinical Sample?	Psychological symptom assessed
McCarthy-Jones & Fernyhough, 2011	No: students	Predisposition to auditory hallucinations (subscale in LSHS-R, (Bentall & Slade, 1985), to visual hallucinations and disturbances (subscale of LSHS-R, (Morrison, Wells, & Nothard, 2000), and anxiety and depression (HADS, (Zigmond & Snaith, 1983).
Alderson-Day et al., 2014	No: students	Experience of auditory hallucinations (subscale in LSHS-R, (Bentall & Slade, 1985), dissociative traits (DES-II (Carlson & Putnam, 1993) and self-esteem (RSES, (Rosenberg, 1965).
Alderson-Day & Fernyhough, 2015	No: students	When experience sampling via smartphone app recorded happiness and anxiety on visual analogue scale from 0 – 10.
Payne & Manning, 1991	No: Pre-training teachers	Self-esteem (self-esteem inventory (Coopersmith, 1967) and locus of control (Locus of Control scale for Teachers (Sadowski, 1982)
Langdon et al., 2009	Yes: patients with schizophrenia	Experience of auditory verbal hallucinations (SAPS item 1, (Andreasen, 1984).

Notes: LSHS-R - Launay-Slade Hallucination Scale Revised. HADS - Hospital Anxiety and Depression Scale. DES-II – Dissociative Experiences Scale – Second Revision. RSES – Rosenberg Self-Esteem Scale. SAPS – Scale for Assessment of Positive Symptoms.

One of the studies in this review measured inner speech differences in a clinical sample. Langdon et al. (2009) found that individuals with a diagnosis of schizophrenia had greater variability in the levels of inner speech that they engaged in, than controls did. Specifically, patients tended to report that they engaged in inner speech very frequently, or hardly at all (Langdon et al., 2009).

These preliminary findings suggest that certain variations in the content, frequency and structure of individual's inner speech may be related to psychological disorders and symptoms.

Discussion

This systematic review sought to critically appraise methods for measuring individual differences in inner speech. It also investigated whether studies that employed such methods found evidence to suggest that inner speech variation is related to psychopathology.

Methods of measuring inner speech

It is possible to measure inner speech as a process in its own right, with questionnaire and interview, introspective, experimental, neuroimaging and lesion methodologies. Table 10 summarises these methods, along with the qualities of the inner speech which the measurement method relies on, and the characteristics of inner speech that are assessed.

The VISQ (McCarthy-Jones & Fernyhough, 2011) was found to have very good measurement properties according to the COSMIN checklist (Terwee et al., 2012). The RSQ (Delamillieure et al., 2010), had much poorer properties according to this checklist, and it was not possible to assess many of the psychometric properties of the interview developed by Langdon et al. (2009). The current best questionnaire and interview measure of inner speech is therefore seen to be the VISQ. Questionnaires are not able to capture spontaneous inner speech as it occurs, but they are presumably asking individuals to reflect on their own retrospective spontaneous inner speech in a general manner.

Introspective methods produce raw data with regard to people's idiosyncratic inner experiences. Researchers are able to develop and apply their own coding categories to this raw data, in order to assess variation in particular phenomena that they may be interested in. For example Payne and Manning (1991) coded the valence, locus of control and transactional states of inner speech. Through data extracted by their coding scheme they found that their sample of training teachers engaged mainly in negative, child oriented and externally focused self-talk (Payne & Manning, 1991). The reliability of these coding categories can be assessed if more than one rater is used. The introspective techniques were found to have good inter-rater reliability if these statistics were reported (Hurlburt & Heavey, 2002; Payne & Manning, 1991).

Table 10. Current measurement techniques and what they can detect

Method	*Name*	*Characteristics of IS*			*Aspects of IS*
		Obj/Sbj	*Rtp/Pres*	*Spon/Elic*	
Questionnaires	VISQ	Sbj	Rtp	NA	Structure, Content
	RSQ	Sbj	Rtp	NA	Frequency, Valence, Content
	'IS interview'	Sbj	Rtp	NA	Frequency, Valence, Content
Introspection	DES	Sbj	Pres	Spon	(researcher interest)
	Thought listing	Sbj	Rtp	NA	(researcher interest)
Experimental	Rhyming judgements	Both	Pres	Elic	Structure
	EMG	Obj	Pres	Elic	Presence
Neuroimaging	fMRI	Both	Pres	Elic	Associated brain areas
	EEG	Both	Pres	Elic	Presence?
Lesion	Various	Obj	Pres	Elic	Associated brain areas, Dissociation from other cognitive abilities

Notes. Obj – objective, Sbj – subjective, Rtp – retrospective, Pres – present moment, Spon – spontaneous, Elic – elicited.

DES was the only introspective method that did not have an a-priori focus on a particular aspect of inner experience, as listing of one's inner thoughts obviously requires individuals to focus on one aspect (Morin et al., 2011; Payne & Manning, 1991). In addition, the

introspective method DES was the only technique of those critiqued in this review that can measure spontaneous inner speech. Hurlburt et al. (2013, p. 1479) explicitly state that they developed DES in order to elicit information about inner speech which was "not expressly manipulated by some experimental procedure" and was "undisturbed by the intention to apprehend it".

Experimental techniques provide some objective methods of assessing the structure of inner speech through silent rhyming judgement tasks (e.g. Lyxell et al., 1994), as well as the possibility of observing whether inner speech is actually occurring by using EMG (e.g. Livesay et al., 1996). Currently, these methods have only been used to investigate individual's elicited inner speech. In addition, as far as is known, these methods have not yet been used to investigate the relationship between inner speech variation and particular psychological disorders and symptoms.

Neuroimaging methods allow for the objective assessment of brain areas that are associated with inner speech activity, although it is not possible to determine objectively whether participants are engaging in inner speech during these tasks. All studies in this review used tasks that elicited participants' inner speech, with the exception of Kuhn et al., (2014) (see below). Elicited inner speech has been described as "an unnatural phenomenon (Jones & Fernyhough, 2007), occurring very rarely outside of psychological research laboratories" (Kühn et al., 2014, p. 6). Lesion studies also provide objective assessment of the relationship between inner speech and other cognitive abilities, as well as information about brain areas that are needed to support inner speech.

Two studies which combined inner speech measurement methods were also critiqued by this review. Kühn et al. (2014) combined DES with fMRI scans, and compared this method with traditional fMRI tasks that rely on elicited inner speech. This proof-of-principal study indicated that DES classified inner speech activates the inferior frontal gyrus, which has been associated with inner speech in previous fMRI studies of elicited inner speech (e.g. Abe et al., 2011). This finding appears to validate both fMRI tasks and the DES method. Although this study presented the results of only one participant, the results indicated that elicited inner speech activates a broader network of brain areas than DES classified spontaneous inner speech (Kühn et al., 2014). Kühn et al. (2014) suggest that eliciting inner speech may result in additional processing, of the prompt itself for example.

Alderson-Day and Fernyhough (2015) found that measuring inner speech by questionnaire self-report tended to result in greater levels of endorsement of various characteristics, than when information was gathered through randomly sampled moments. This provides support for Hurlburt et al.'s (2013) argument that asking about inner speech in generalised, retrospective forms may lead to inflated responses. In addition, the size of the discrepancy between questionnaire and sampling report varied, depending on the inner speech characteristic that was being measured (Alderson-Day & Fernyhough, 2015). Specifically, reliability between questionnaire and sampling reports of evaluative and motivational inner speech was divergent. Alderson-Day and Fernyhough (2015) tentatively suggested that this sort of inner speech may be linked to negative beliefs and ideas about the self, which may be more salient than other phenomenological properties of inner speech when thinking in general and retrospective terms.

Individual differences in inner speech and other experiences

This review demonstrates that these methodologies can be used to discern individual differences in inner speech. There are preliminary suggestions that these measurement methods provide data about inner speech which is related to individual differences in other human experiences. For example, the frequency of engaging in inner speech without other inner experiences, was inversely related to the speed of people's out-loud speech (Hurlburt et al., 2002). There is also a suggestion that there may be relationships between demographics and inner speech differences. Delamillieure et al. (2010) found that a higher education level was associated with individuals reporting greater content related to the future in their inner speech.

Relationships between inner speech and psychopathology

This review demonstrates that there is preliminary evidence to suggest that inner speech differences are related to psychological disorders and symptoms. For example, the amount of inner speech which individuals engaged in was found to be different if someone did or did not have a diagnosis of schizophrenia (Langdon et al., 2009). In addition, the tendency of individuals (although from a non-clinical sample), to experience auditory verbal hallucinations was also found to relate to their experience of the presence of others' voices in their inner

speech and to evaluative/motivational and dialogic inner speech (Alderson-Day et al., 2014; McCarthy-Jones & Fernyhough, 2011).

The content of inner speech also appears to have relationships with psychological symptoms (at least in non-clinical samples). Self-esteem was inversely related to the presence of other people's voices in one's inner speech as measured by questionnaires (Alderson-Day et al., 2014). Anxiety levels were related to both questionnaire and momentary sampling measures of the presence of other's voices in inner speech (Alderson-Day & Fernyhough, 2015; McCarthy-Jones & Fernyhough, 2011). Using inner speech to evaluate and motivate oneself was also associated with anxiety when measured by questionnaire methods (McCarthy-Jones & Fernyhough, 2011). Alderson-Day et al. (2014) note that the VISQ, which was used to discern these relationships, does not elicit the valence related to any of the characteristics measured. For example some evaluations could be "more positive (e.g. "I *can* do this") and others negative ("I really shouldn't have done that")" (Alderson-Day et al., 2014, p. 294). Valence is also likely to be associated with the presence of other peoples' voices (e.g. critical versus supportive). It is possible that the relationships between inner speech and anxiety or self-esteem may be related to this content having a negative valence.

However, the relationship between using one's evaluative inner speech and self-reported anxiety was not found when inner speech was measured by experience sampling (Alderson-Day & Fernyhough, 2015). As discussed above, Alderson-Day and Fernyhough (2015) suggest that this may be related to the salience of particular phenomenological properties of inner speech, when thinking about oneself in general terms.

The literature investigated in this review offers preliminary suggestions that inner speech variation may be related to individual differences in psychological disorders and symptoms. It is suggested that the presence of such relationships warrant further investigation.

Limitations and future research

This review has highlighted that although it is possible to measure individual differences in inner speech, the methodologies that are currently available require further development. The VISQ was only recently developed, and requires replication in samples outside a university student populations. In addition, all of the questionnaire and interview, and

introspective measures critiqued here, were administered by the authors who developed them, and therefore their findings also require replication by different laboratories (Heavey & Hurlburt, 2008).

Questionnaire methods could be developed in order to assess additional aspects of inner speech. As noted by the authors themselves (McCarthy-Jones & Fernyhough, 2011), the VISQ in its current form does not capture information about whether individuals experience affect along with their inner speech utterances, as is suggested by Delamillieure et al.'s (2010) findings. Nor does the VISQ ask how much individuals use inner speech to self-regulate, or to support their memory (McCarthy-Jones & Fernyhough, 2011). These aspects are measured by some of the questionnaires that have been developed to investigate self-directed talk more broadly, with the SVQ (Duncan & Cheyne, 1999) assessing overt self-regulation and mnemonic functions for example (McCarthy-Jones & Fernyhough, 2011). However there are currently no questionnaires that measure these functions specifically in inner speech.

Alderson-Day and Fernyhough's (2015) study indicates that it is possible to gather subjective experience sampling introspective data via a smartphone app, in a much less labour intensive way than is required by DES for both participants and researchers. Kühn et al.'s (2014) proof-of-principal investigation shows that it is possible to combine objective methods of measurement such as fMRI, with subjective methodologies such as DES. It would be interesting to continue to attempt to corroborate individuals' subjective reporting of engagement in inner speech (through DES or an experience sampling app for example) with objective measure (using EMG, to investigate covert lip muscle activity associated with inner speech for example). In an EMG experiment investigating a different form of inner experience, Rapin, Dohen, Polosan, Perrier, and Lœvenbruck (2013), found an increase in covert lip muscle activity during auditory verbal hallucinations. Such corroborative studies are important for validating methods which individually can only measure inner speech either objectively or subjectively.

The majority of the papers included in this review measured inner speech with just one methodology. The two papers which combined measures validated aspects of the methods of measurement. For example, it appears that inner speech as categorised by DES is associated with activation in some of the brain areas that have been found to be associated with elicited inner speech during fMRI scans. However this study also indicated that some of the brain

areas associated with elicited inner speech may not be associated with spontaneous inner speech, although this finding needs replication in other studies with greater numbers.

Alderson-Day and Fernyhough's (2015) corroborative study has also shown that it is possible to compare retrospective questionnaire reports with in-the-moment experience sampling, and thereby provide validation for aspects of these measurement methods. That is, some characteristics of inner speech (such as its condensed or extended nature) can be assessed by both questionnaire and momentary sampling methods, but it may not be valid or reliable to measure other characteristics (such as how much people use inner speech to evaluate and motivate) through questionnaire methods. Therefore it appears that the reliability and validity of participants subjective and retrospective reporting of their own spontaneous inner speech may depend on the particular characteristic of inner speech being investigated.

These two studies highlight the importance for future studies that investigate measurement of inner speech to combine modalities of measurement. Combining allows validation of the different methods. It also indciates whether findings about individual differences in inner speech are valid and reliable and are not instead an artefact of the method of measurement used.

Conclusion

The methods that exist for measuring inner speech vary with regard to the characteristics which they are able to measure, such as its frequency, content or structure. Methods also vary regarding the qualities of the inner speech which they rely on, such as whether it is subjective, retrospective or elicited. Combining methods of measurement enables the validation of methods that rely on differing qualities of inner speech. There are preliminary findings that individual differences in inner speech found by these methods of measurement, are related to variation in other aspects of human experience. This includes variation in psychological disorders and symptoms such as auditory verbal hallucinations, anxiety and self-esteem. It is suggested that a more thorough assessment of inner speech processes across psychological disorders would be useful for the advancement of clinical psychology theory and treatment (Pearson et al., 2013).

References

Abe, K., Takahashi, T., Takikawa, Y., Arai, H., & Kitazawa, S. (2011). Applying independent component analysis to detect silent speech in magnetic resonance imaging signals. *European Journal of Neuroscience, 34*(8), 1189-1199.

Alderson-Day,B., & Fernyhough,C. (in press). Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology.

Alderson-Day, B., & Fernyhough, C. (2015). Relations among questionnaire and experience sampling measures of inner speech: A smartphone app study. *Frontiers in psychology, 6*, 517.

Alderson-Day, B., McCarthy-Jones, S., Bedford, S., Collins, H., Dunne, H., Rooke, C., & Fernyhough, C. (2014). Shot through with voices: Dissociation mediates the relationship between varieties of inner speech and auditory hallucination proneness. *Consciousness and cognition, 27*, 288-296.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental neuropsychology, 20*(1), 385-406.

Andreasen, N. C. (1984). Scale for the assessment of positive symptoms. *Iowa City: University of Iowa*.

Baars, B. (2003). How brain reveals mind neural studies support the fundamental role of conscious experience. *Journal of Consciousness Studies, 10*(9-10), 100-114.

Baciu, M. V., Rubin, C., Décorps, M. A., & Segebarth, C. M. (1999). fMRI assessment of hemispheric language dominance using a simple inner speech paradigm. *NMR in Biomedicine, 12*(5), 293-298.

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: evidence from task switching. *Journal of Experimental Psychology: General, 130*(4), 641.

Behar, E., Zuellig, A. R., & Borkovec, T. (2005). Thought and imaginal activity during worry and trauma recall. *Behavior Therapy, 36(2),* 157-168.

Bentall, R., & Slade, P. D. (1985). Reality testing and auditory hallucinations: a signal detection analysis. *British journal of clinical psychology, 24*(3), 159-169.

Berk, L. E. (1986). Relationship of elementary school children's private speech to behavioral accompaniment to task, attention, and task performance. *Developmental Psychology, 22*(5), 671.

Berne, E. (1968). *Games people play: The psychology of human relationships* (Vol. 2768): Penguin UK.

Brinthaupt, T. M., Hein, M. B., & Kramer, T. E. (2009). The self-talk scale: Development, factor analysis, and validation. *Journal of personality assessment, 91*(1), 82-92.

Burnett, P. C. (1996). Children's Self-talk and Significant Others' Positive and Negative Statements. *Educational Psychology, 16*(1), 57-67.

Carlson, E. B., & Putnam, F. W. (1993). An update on the dissociative experiences scale. *Dissociation: Progress in the Dissociative Disorders*.

Chang, Y.-K., Ho, L.-A., Lu, F. J.-H., Ou, C.-C., Song, T.-F., & Gill, D. L. (2014). Self-talk and softball performance: The role of self-talk nature, motor task characteristics, and self-efficacy in novice softball players. *Psychology of Sport and Exercise, 15*(1), 139-145.

Coopersmith, S. (1967). *The antecedents of self-esteem*: Consulting Psychologists Pr.

Damianova, M. K., Lucas, M., & Sullivan, G. B. (2012). Verbal mediation of problem solving in pre-primary and junior primary school children. *South African Journal of Psychology, 42*(3), 445-455.

Daugherty, M., White, C. S., & Manning, B. H. (1994). Relationships among private speech and creativity measurements of young children. *Gifted Child Quarterly, 38*(1), 21-26.

Day, K. L., & Smith, C. L. (2013). Understanding the role of private speech in children's emotion regulation. *Early Childhood Research Quarterly, 28*(2), 405-414.

Delamillieure, P., Doucet, G., Mazoyer, B., Turbelin, M.-R., Delcroix, N., Mellet, E., . . . Tzourio-Mazoyer, N. (2010). The resting state questionnaire: an introspective questionnaire for evaluation of inner experience during the conscious resting state. *Brain research bulletin, 81*(6), 565-573.

Duncan, R. M., & Cheyne, J. A. (1999). Incidence and functions of self-reported private speech in young adults: A self-verbalization questionnaire. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 31*(2), 133.

Duncan, R. M., & Cheyne, J. A. (2002). Private speech in young adults: Task difficulty, self-regulation, and psychological predication. *Cognitive development, 16*(2002), 889-906.

Fernyhough, C. (2004). Alien voices and inner dialogue: towards a developmental account of auditory verbal hallucinations. *New ideas in Psychology, 22*(1), 49-68.

Ford, J. M., Mathalon, D. H., Kalba, S., Whitfield, S., Faustman, W. O., & Roth, W. T. (2014). Cortical responsiveness during inner speech in schizophrenia: an event-related potential study.

Geva, S., Bennett, S., Warburton, E. A., & Patterson, K. (2011). Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. *Aphasiology, 25*(3), 323-343.

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J.-C., & Warburton, E. A. (2011). The neural correlates of inner speech defined by voxel-based lesion–symptom mapping. *Brain, 134*(10), 3071-3082.

Hardy, J., Hall, C. R., & Hardy, L. (2005). Quantifying athlete self-talk. *Journal of Sports Sciences, 23*(9), 905-917.

Hatzigeorgiadis, A., Zourbanos, N., Galanis, E., & Theodorakis, Y. (2011). Self-Talk and Sports Performance A Meta-Analysis. *Perspectives on Psychological Science, 6*(4), 348-356.

Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness and cognition, 17*(3), 798-810.

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *bmj, 343*.

Hinke, R. M., Hu, X., Stillman, A. E., Kim, S.-G., Merkle, H., Salmi, R., & Ugurbil, K. (1993). Functional magnetic resonance imaging of Broca's area during internal speech. *Neuroreport, 4*(6), 675-678.

Holmes, E. A., Grey, N., & Young, K. A. (2005). Intrusive images and "hotspots" of trauma memories in posttraumatic stress disorder: An exploratory investigation of emotions

and cognitive themes. *Journal of behavior therapy and experimental psychiatry, 36*(1), 3-17.

Holmes, E. A., Mathews, A., Dalgleish, T., & Mackintosh, B. (2006). Positive interpretation training: Effects of mental imagery versus verbal training on positive mood. *Behavior Therapy, 37(3),* 237-247.

Hurlburt, R. T., & Heavey, C. L. (2002). Interobserver reliability of descriptive experience sampling. *Cognitive Therapy and Research, 26*(1), 135-142.

Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and cognition, 22*(4), 1477-1494.

Hurlburt, R. T., Koch, M., & Heavey, C. L. (2002). Descriptive experience sampling demonstrates the connection of thinking to externally observable behavior. *Cognitive Therapy and Research, 26*(1), 117-134.

Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and cognition, 16*(2), 391-399.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology review, 17*(3), 213-233.

Kay, J., Coltheart, M., & Lesser, R. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia,(PALPA): Auditory Processing*: Lawrence Erlbaum Associates.

Kendall, P. C., Howard, B. L., & Hays, R. C. (1989). Self-referent speech and psychopathology: The balance of positive and negative thinking. *Cognitive Therapy and Research, 13*(6), 583-598.

Korba, R. J. (1990). The rate of inner speech. *Perceptual and motor skills, 71*(3), 1043-1052.

Kühn, S., Fernyhough, C., Alderson-Day, B., & Hurlburt, R. T. (2014). Inner experience in the scanner: can high fidelity apprehensions of inner experience be integrated with fMRI? *Frontiers in psychology, 5*.

Langdon, R., Jones, S., Connaughton, E., & Fernyhough, C. (2009). The phenomenology of inner speech: comparison of schizophrenia patients with auditory verbal hallucinations and healthy controls. *Psychological medicine, 39*(04), 655-663.

Lee, J.-H., Marzelli, M., Jolesz, F. A., & Yoo, S.-S. (2009). Automated classification of fMRI data employing trial-based imagery tasks. *Medical image analysis, 13*(3), 392-404.

Levine, D. N., Calvanio, R., & Popovics, A. (1982). Language in the absence of inner speech. *Neuropsychologia, 20*(4), 391-409.

Lidstone, J. S., Meins, E., & Fernyhough, C. (2010). The roles of private speech and inner speech in planning during middle childhood: Evidence from a dual task paradigm. *Journal of experimental child psychology, 107*(4), 438-451.

Livesay, J., Liebke, A., Samaras, M., & Stanley, A. (1996). Covert speech behavior during a silent language recitation task. *Perceptual and motor skills, 83*(3f), 1355-1362.

Luria, A. R. (1965). LS Vygotsky and the problem of localization of functions. *Neuropsychologia, 3*(4), 387-392.

Lyxell, B., Rönnberg, J., & Samuclsson, S. (1994). Internal speech functioning and speechreading in deafened and normal hearing adults. *Scandinavian audiology, 23*(3), 179-185.

McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and cognition, 20*(4), 1586-1593.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151*(4), 264-269.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology, 63*(7), 737-745.

Morin, A., Uttl, B., & Hamper, B. (2011). Self-reported frequency, content, and functions of inner speech. *Procedia-Social and Behavioral Sciences, 30*, 1714-1718.

Morrison, A. P., Wells, A., & Nothard, S. (2000). Cognitive factors in predisposition to auditory and visual hallucinations. *British journal of clinical psychology, 39*(1), 67-78.

Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition, 106*(1), 528-537.

Owens, D. K., Lohr, K. N., Atkins, D., Treadwell, J. R., Reston, J. T., Bass, E. B., . . . Helfand, M. (2010). AHRQ series paper 5: grading the strength of a body of evidence when

comparing medical interventions—Agency for Healthcare Research and Quality and the Effective Health-Care Program. *Journal of clinical epidemiology, 63*(5), 513-523.

Patel, T., Brewin, C. R., Wheatley, J., Wells, A., Fisher, P., & Myers, S. (2007). Intrusive images and memories in major depression. *Behaviour Research and Therapy, 45*(11), 2573-2580.

Payne, B. D., & Manning, B. H. (1991). Self-talk of student teachers and resulting relationships. *The Journal of Educational Research, 85*(1), 47-51.

Pearson, D. G., Deeprose, C., Wallace-Hadrill, S., Heyes, S. B., & Holmes, E. A. (2013). Assessing mental imagery in clinical psychology: A review of imagery measures and a guiding framework. *Clinical psychology review, 33*(1), 1-23.

Rapin, L., Dohen, M., Polosan, M., Perrier, P., & Lœvenbruck, H. (2013). An EMG study of the lip muscles during covert auditory verbal hallucinations in schizophrenia. *Journal of Speech, Language, and Hearing Research, 56*(6), S1882-S1893.

Rapin, L. A., Dohen, M., Loevenbruck, H., Whitman, J. C., Metzak, P. D., & Woodward, T. S. (2012). Hyperintensity of functional networks involving voice-selective cortical regions during silent thought in schizophrenia. *Psychiatry Research: Neuroimaging, 202*(2), 110-117.

Rosenberg, M. (1965). *Society and the adolescent self-image*: Princeton University Press Princeton, NJ.

Sadowski, C. J. (1982). *The reliability and validity of a Likert-type locus of control scale for teachers*: American Psycholog. Ass., Journal Suppl. Abstract Service.

Sanderson, S., Tatt, I. D., & Higgins, J. P. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *International journal of epidemiology, 36*(3), 666-676.

Schneider, J. F., Pospeschill, M., & Ranger, J. (2005). Does self-consciousness mediate the relation between self-talk and self-knowledge? . *Psychological Reports, 96*(2), 387-396.

Scholkmann, F., Klein, S. D., Gerber, U., Wolf, M., & Wolf, U. (2014). Cerebral hemodynamic and oxygenation changes induced by inner and heard speech: a study combining functional near-infrared spectroscopy and capnography. *Journal of biomedical optics, 17002*, 1.

Scholkmann, F., Wolf, M., & Wolf, U. (2013). The effect of inner speech on arterial CO2 and cerebral hemodynamics and oxygenation: a functional NIRS study *Oxygen Transport to Tissue XXXV* (pp. 81-87): Springer.

Shergill, S. S., Brammer, M. J., Fukuda, R., Williams, S. C., Murray, R. M., & McGUIRE, P. K. (2003). Engagement of brain areas implicated in processing inner speech in people with auditory hallucinations. *The British Journal of Psychiatry, 182*(6), 525-531.

Siegrist, M. (1995). Inner speech as a cognitive process mediating self-consciousness and inhibiting self-deception. *Psychological Reports, 76*(1), 259-265.

Simons, C. J., Tracy, D. K., Sanghera, K. K., O'Daly, O., Gilleen, J., Krabbendam, L., & Shergill, S. S. (2010). Functional magnetic resonance imaging of inner speech in schizophrenia. *Biological psychiatry, 67*(3), 232-237.

Stokes, C., & Hirsch, C. R. (2010). Engaging in imagery versus verbal processing of worry: Impact on negative intrusions in high worriers. *Behaviour Research and Therapy, 48(5),* 418-423.

Taylor, P., Hutton, P., & Dudley, R. (2014). Rationale and protocol for a systematic review and meta-analysis on reduced data gathering in people with delusions. *Systematic reviews, 3*(1), 44.

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W., Bouter, L. M., & de Vet, H. C. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research, 21*(4), 651-657.

Uttl, B., Morin, A., & Hamper, B. (2011). Are inner speech self-report questionnaires reliable and valid? *Procedia-Social and Behavioral Sciences, 30*, 1719-1723.

Vercammen, A., Knegtering, H., Bruggeman, R., & Aleman, A. (2011). Subjective loudness and reality of auditory verbal hallucinations and activation of the inner speech processing network. *Schizophrenia bulletin, 37*(5), 1009-1016.

Vicente, A., & Martinez Manrique, F. (2011). Inner speech: Nature and functions. *Philosophy Compass, 6*(3), 209-219.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: The MIT Press.

Winsler, A. (2009). Still talking to ourselves after all these years: A review of current research on private speech. *Private speech, executive functioning, and the development of verbal self-regulation*, 3-41.

Winsler, A., De Leon, J. R., Wallace, B. A., Carlton, M. P., & Willson-Quayle, A. (2003). Private speech in preschool children: Developmental stability and change, across-task consistency, and relations with classroom behaviour. *Journal of Child Language, 30*(03), 583-608.

Winsler, A., & Naglieri, J. (2003). Overt and covert verbal problem-solving strategies: Developmental trends in use, awareness, and relations with task performance in children aged 5 to 17. *Child development, 74*(3), 659-678.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica, 67*(6), 361-370.

Appendices

Appendix 1: PsycInfo search Strategy

1. self-talk.mp. or exp Self Talk/

2. "inner speech".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

3. "internal speech".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

4. "internal dialogue".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

5. "inner dialogue".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

6. "internal monologue".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

7. "inner monologue".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

8. "self-directed speech".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

9. "verbal mediation".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

10. "private speech".mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

11. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10

12. measur*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

13. exp Measurement/

14. exp Psychometrics/

15. psychometric*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

16. assess*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

17. questionnaire*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

18. exp Questionnaires/

19. exp Introspection/

20. introspect*.mp.

21. quantif*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

22. examinat*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

23. evaluat*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

24. exp Evaluation/

25. investigat*.mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]

26. 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25

27. 11 and 26

28. limit 27 to (human and english language)

Appendix 2: Medline search Strategy

1. "self-talk".mp.

2. "inner speech".mp.

3. "internal speech".mp.

4. "internal dialogue".mp.

5. "inner dialogue".mp.

6. "internal monologue".mp.

7. "inner monologue".mp.

8. "self-directed speech".mp.

9. "verbal mediation".mp.

10. "private speech".mp.

11. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10

12. measur*.mp.

13. Psychometrics/

14. psychometric*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

15. assess*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

16. questionnaire*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

17. Questionnaires/

18. introspect*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

19. quantif*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

20. examinat*.mp.

21. evaluat*.mp.

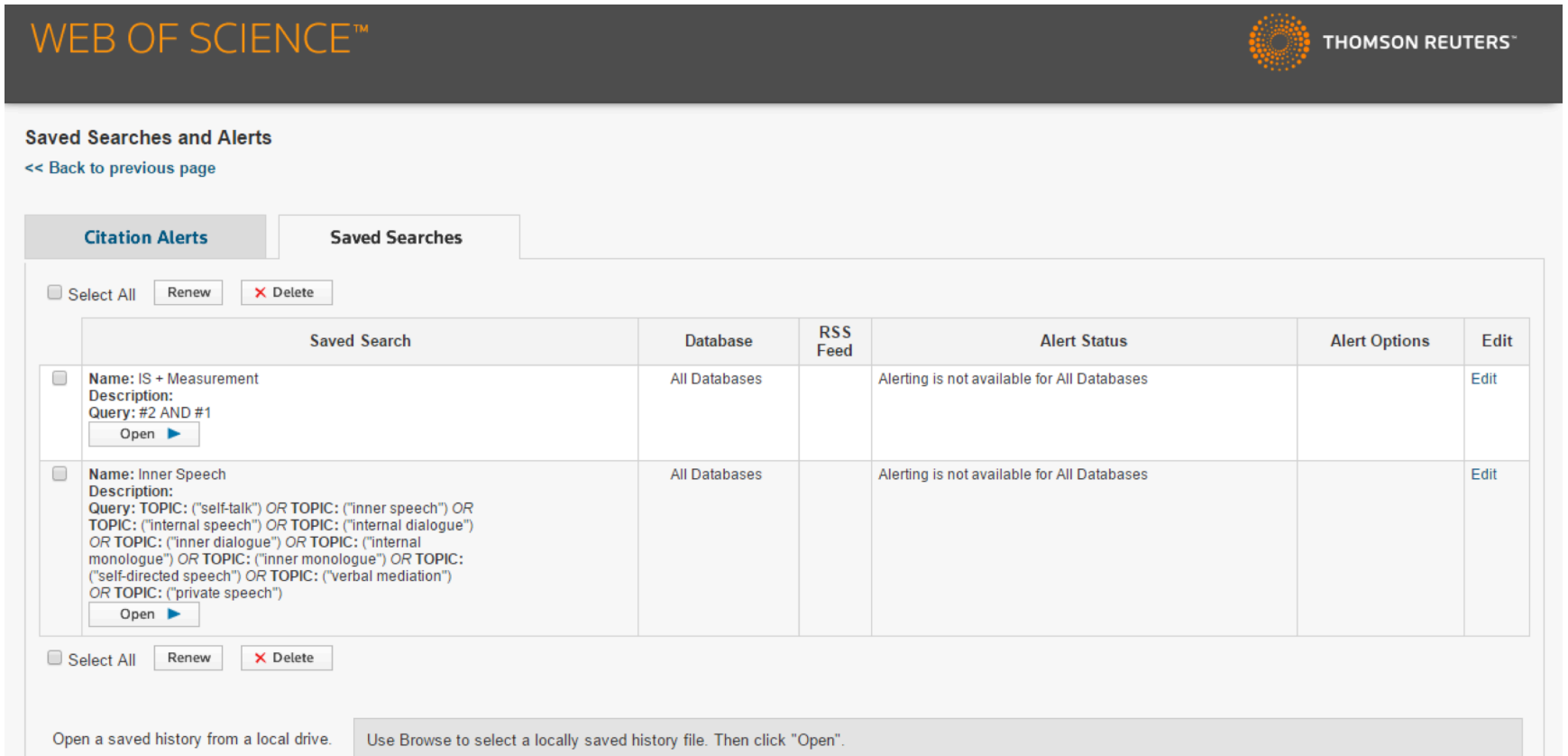
22. investigat*.mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

23. 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22

24. 11 and 23

25. limit 24 to (english language and humans)

Appendix 3: Web of Science search strategy



WEB OF SCIENCE™

THOMSON REUTERS

Saved Searches and Alerts

<< Back to previous page

Citation Alerts | Saved Searches

Select All | Renew | Delete

	Saved Search	Database	RSS Feed	Alert Status	Alert Options	Edit
☐	**Name:** IS + Measurement **Description:** **Query:** #2 AND #1 Open ▶	All Databases		Alerting is not available for All Databases		Edit
☐	**Name:** Inner Speech **Description:** **Query:** **TOPIC:** ("self-talk") *OR* **TOPIC:** ("inner speech") *OR* **TOPIC:** ("internal speech") *OR* **TOPIC:** ("internal dialogue") *OR* **TOPIC:** ("inner dialogue") *OR* **TOPIC:** ("internal monologue") *OR* **TOPIC:** ("inner monologue") *OR* **TOPIC:** ("self-directed speech") *OR* **TOPIC:** ("verbal mediation") *OR* **TOPIC:** ("private speech") Open ▶	All Databases		Alerting is not available for All Databases		Edit

Select All | Renew | Delete

Open a saved history from a local drive. | Use Browse to select a locally saved history file. Then click "Open".

Appendix 4: Sample characteristics, inclusion/exclusion criteria and selection procedure

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
McCarthy-Jones & Fernyhough, 2011	2 samples of university students	235	18-30 (20.4, 2.9).	77 M, 158 F.	Not reported	Assume British English	Online	None reported	Recruited by e-mail invitation, received no financial incentive	NA
		220	18-30 (23, 3.5).	47 M, 173 F.						
Alderson-Day et al., 2014	University students	156	18-31 (20, 2.1).	22 M, 134 F.	Not reported	Assume British English	University	None reported	Web advert and course credit provided	NA
Delamillieure et al., 2010	Healthy volunteers	180	18 – 51 (26, 6.6).	89 M, 91 F.	Not reported	Assume French	Not reported	No history of any psychiatric, neurological or medical conditions	Not reported	NA

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Langdon et al., 2009	Patients with schizophrenia	29	-- (41.2, 10.9)	15 M, 14 F	Not reported	Good English skills	Outpatient clinics	Exclusion - prominent thought disorder, current substance abuse, mental retardation, clinically significant head injury.	Not reported	Mean age onset 25.2 (SD = 7.6). Mean illness duration 15.8, (SD = 8.8). Mean AVHs (SAPS item 1) = 2.6 (SD = 2.2).
	Controls	42	-- (36.8, 13.1)	24 M, 18 F			General community	Inclusion - match patient group on age, sex, IQ.		

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
(Hurlburt & Heavey, 2002)	University students	10	18 – 19 (18.3, --)	5 M, 5 F	Not reported	Not reported	University	Not reported	210 students completed SCL-90-R and divided into deciles based on score. Randomly selected 1 from each decile. If declined participation, another selected: 3 x schedule conflict, 1 x lack of interest.	NA

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Hurlburt & Heavey, 2002	University students: High-speech rate	7	18-25	2 M, 5 F.	Not reported	6 American English native speakers, 1 native language German, but fluent English.	Students from introductory psychology university course	Not reported	HIGH SPEED 17 of 348 students endorsed talking 'a lot faster' than others. 13 of these (for course credit) had objective measure of speech rate. 8 of these had high speech rate. 7 of these participated.	NA
	Comparison group.	7	18 - 21	2 M, 5 F.		Not reported – assume American English			7 students matched for sex and approximate age.	

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Heavey & Hurlburt, 2008	University students	30	-- (19, 3.3)	14 M, 16 F.	63% Caucasian, 17% Asian, 10% Hispanic, 7% African American, 3% Native American	Assume American English	University	None reported	407 students completed SCL-90-R and divided into deciles based on score. Randomly selected 3 from each decile.	NA
Morin et al., 2011	University students	380	17 - 55 (21.3, --)	65 M, 315 F	Not reported	89% declared English as first language	University	None reported	"Volunteers"	NA
Payne & Manning, 1991	Pre-service teachers	69	Not reported	69 F	100% White	Assume English	Not clear	None reported	None reported	NA
Lyxell et al., 1994	Deafened adults	15	-- (53, --)	8 M, 7 F	Not reported	Assume Swedish	Not reported	Not reported	Not reported	deafened adults had no functional residual hearing
	Controls	22	-- (53, --)	9 M, 11 F						

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Oppenheim & Dell, 2008	Local residents	48	20-30	Not reported	Not reported	American English. Had not learnt another language before 5 years.	Residents from the local town	Normal or corrected-to-normal vision and hearing	Not reported. Participants received $10.	NA
Korba, 1990	University students	38	-- (19.2, --)	17 M, 21 F	Not reported	Assume English	University	Not reported	'randomly selected' from course.	NA
Livesay et al., 1996	Volunteers from a university setting	20	-- (38.7, 12.3)	8 M, 12 F.	Not reported	Assume English	Private research school	Not reported	Not reported	NA
Hinke et al., 1993	Healthy participants	9	21 - 40	4 M, 5 F	Not reported	Assume English	Academic environment	Not reported	Not reported	NA
Baciu et al., 1999	Healthy participants	10	-- (37, --)	10 M	Not reported	Assume French	Not clear	Not reported	Not reported	NA

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Lee et al., 2009	Healthy participants	5	22 - 35	3 M, 2 F	Not reported	Assume American English	Not clear	Non-smokers, no history of neurological or psychiatric conditions	Not reported	NA
Abe et al., 2011	Healthy participants	8	20 - 29	1 M, 7 F	Not reported	Assume Japanese	Not clear	No history of neurological disease	Not reported	NA
Scholkmann et al., 2013	Healthy participants	7	-- (34.6, 9.3)	4 M, 3 F	Not reported	German/Swiss German native speakers	Not clear	Not reported	Not reported	NA
Scholkmann et al., 2014	Healthy participants	29	-- (47, 12.8)	14 M, 15 F	Not reported	German/Swiss German native speakers	Not clear	Not reported	Not reported	NA

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Shergill et al., 2003	Patients with SZ	8	20 - 45 (31, 9)	8 M	Not reported	Assume English	Not clear	No history of head injury, neurological symptoms, speech or hearing difficulties, did not meet DSM-IV criteria of misuse /dependence on drugs/alcohol during lifetime.	Patients recruited from wards and clinics at Maudsley Hospital	DSM-IV diagnosis of schizophrenia. Had experienced AVHs.
	Controls	8(?)	23 - 37 (29, 5)	8 M				No medical/psychiatric disorder, medication, family history of psychiatric disorder.	Not reported	
Simons et al., 2010	Patients with SZ	15	-- (34.7, 8.7)	15 M	Not reported	First language English	Not clear	No drug use within previous 6 months.	Recruited through consultant and key worker recommendations	DSM-IV diagnosis of schizophrenia. Had experienced AVHs. Taking antipsychotic medication. PANS mean score 48.5 (SD = 16.5). Mean illness 11.2 years.
	Controls	12	-- (34.4, 7.9)	15 M				No psychiatric disorders or family history of psychiatric disorder	City-wide newspaper advert	

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Vercammen et al., 2011	Patients with SZ	22	19 - 60 (36.2, 12.3)	11 M, 11 F	Not reported	Assume Dutch	Not clear	Not reported	Recruited from inpatient and outpatient facilities.	DSM-IV diagnosis of schizophrenia. 21 on antipsychotic medication. All medication resistant with regard to AVHs. Mean illness duration (years) 13 years, (SD =12.6).

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Rapin et al., 2012	Patients with SZ	5	-- (33.6, 7.8)	3 M, 2 F	Not reported	Had been using English daily for 5 years	Not clear	No strong negative symptoms or severe thought disorder. Both groups no history of neurological disorder, TBI (with loss of consciousness for more than 5 mins), drug/alcohol abuse/addiction. Also no history of psychiatric disorder (self or close family).	Recruited from General Hospital wards	4 DSM-IV diagnosis of schizophrenia, 1 schizoaffective disorder. All taking atypical neuroleptics. Mean duration of illness 8.2 years (SD = 4.4)
	Controls	10	-- (26.7, 7.5)	5 M, 5 F					Volunteers from adverts	
Ford et al., 2014	Patients with SZ	15	22-53	13 M, 2 F	Not reported	Assume American English	Not clear	No DSM-IV diagnoses according to SCID.	Not reported	DSM-IV diagnosis of schizophrenia
	Controls	15	20-58	13 M, 2 F						

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Levine et al., 1982	Individual who lost ability to subjectively speak silently following stroke	1	54	M	Not reported	Assume English	Not clear	Not reported	Not reported	NA
Geva, et al., 2011	Patients with aphasia	21	21 - 81 (64, 15)	14 males, 7 females	Not reported	Native speakers of English	Not clear	Aphasia after stroke, no history of neurological or psychiatric disorders (other than stroke) and no major cognitive impairment	Not reported	Left middle cerebral artery territory stroke. Aphasia diagnosed by standardised examination.

Author, year	Sample Overview	N	Age range (M, SD)	Sex	Ethnicity	Language	Settings	Inclusion/Exclusion adequate/reported?	Selection procedure	Disease Characteristics
Geva et al., 2011	Patients with aphasia	29	21 – 81 (64.6, 13.4)	20 M, 9 F	Not reported	Native speakers of British English	Not clear	Aphasia after stroke, no history of neurological or psychiatric disorders (other than stroke) and no major cognitive impairment. 6 months after stroke	Not reported	Left middle cerebral artery territory stroke. Diagnosis of aphasia based on standardised examination.
	Controls	27	44 - 72 (62.2, 8.2)	12M, 15F				No previous history of stroke, neurological or psychiatric disorders		
Kühn et al., 2014	Unknown	1	18	F	Not reported	Not reported	Not clear	None reported	Not reported	NA
Alderson-Day & Fernyhough, 2015	University students	51	-- (19.8, 3.0)	13 M, 38 F	Not reported	English as first language	University	None reported	Recruited by participant-pool advertisement.	NA

Notes: -- - Missing Data. AVHs – Auditory Verbal Hallucinations, SAPS – Scales for the Assessment of Positive Symptoms. SCL-90-R – Symptom Checklist 90 Revised. AVHs – Auditory Verbal Hallucinations, SAPS – Scales for the Assessment of Positive Symptoms.

Appendix 5: Information used to inform COSMIN ratings of questionnaires

Information used to inform COSMIN ratings of questionnaires and interview measures			
Author, year	*McCarthy-Jones & Fernyhough, 2011*	*Delamillieure et al., 2010*	*Langdon et al., 2009*
Missing data acceptable? (Less than 20%)	Missing data formed less than 0.5% of the total responses, and such data points were replaced by the mean response value for that item from all other participants.	Not reported	Not reported
Internal consistency	Based on a formative model. Sample size deemed adequate. Factor analysis performed (see structural validity). Cronbach's alpha of condensed = .8; of dialogic = .83; of other people = .88; of evaluative/motivational = .80.	Did not present internal consistency statistics, or unidimensionality analysis.	Did not present internal consistency statistics, or unidimensionality analysis.
Reliability	36 volunteers completed VISQ one month later. Test-retest reliability: condensed =.61; dialogic = .69; other people =.66; evaluative/motivational =.8	Two measurements for entire questionnaire were not taken; time interval was not appropriate. On section that was re-administered - 14 participants changed quotation by more than 10%.	Two measurements for entire interview were not taken.
Measurement error	Administered twice.	Instrument not administered twice	Interview not administered twice.
Content validity	Did have discussions with students and informal piloting and got rid of items if not loading onto factors. Not sure whether comprehensiveness has been evaluated really, and have left out some aspects of variety in inner speech.	Not clear where 5 mental activities were taken from, and no judgements about relevance and comprehensiveness of items. (Appears to have face-validity).	Not clear where questions were taken from. Has face validity
Structural validity	Bartlett's test of sphericity was significant, therefore conclude that there are correlations in the data set that are appropriate for factor analysis. Kaiser-Meyer-Olking of appropriateness of factor analysis (one final 18 items) = .83 (considered meritorious). Ran confirmatory factor analysis on second sample using AMOS 7 and proposed 4 factor solution was not significantly different from the data. Sample size deemed adequate	NA as this instrument appears to be based on a formative model, not a reflective model.	NA as this instrument appears to be based on a formative model, not a reflective model.
Hypotheses-testing	Presented tentative hypotheses, but not direction. Described measurement properties of comparator instruments.	No hypotheses presented	Hypotheses formulated a priori, and expected direction, but not magnitude provided.
Cross-cultrual validity	NA	NA	NA
Criterion validity	No gold standard measure of variety in inner speech exists.	No gold standard measure exists.	No gold standard measure exists.
Responsiveness	No inner speech intervention was put in place in order to measure change in VISQ scores before and afterwards.	Interview was not administered twice.	Interview was not administered twice.
Interpretability	Frequencies of scores reported: percentage of lowest and highest scores reported.	Did not report distribution (did report mean and SD)	Reported highest and lowest scores.

Empirical Project

An investigation into the relationship between adolescent risk taking behaviour and executive functions, impulsivity and inner speech

First Supervisor: Dr Troy Tranah

Second Supervisors: Dr Eamon McCrory and Professor Essi Viding.

Table of Contents

List of Figures and Tables

Abstract

Adolescence is a time of heightened risk-taking, when males in particular engage in increased levels of behaviour such as alcohol, tobacco and drug use and fighting. However many adolescents do not engage in all such behaviours, indicating that there are marked individual differences in risk-taking. Individual differences in executive function, including working memory, inhibitory control and planning ability, as well as impulsivity have been found to predict adolescent risk-taking.

This study hypothesised that the presence of a situation in which risk-taking could occur activates the executive system, which then recruits the inner speech system to support its functioning. It investigated whether individual differences in both executive functioning and inner speech contributed to an individual's level of impulsivity, increasing or reducing the likelihood of their engagement in real-life risk-taking behaviour.

Sixty 13 to 15 year old male adolescents, recruited from pupil referral units and comprehensive schools, took part in this study. Participants completed measures of executive function, inner speech, cognitive ability impulsivity and real-life risk-taking. Stronger executive function was found to predict reduced engagement in real-world risk-taking behaviour, and this relationship was fully mediated by self-reported impulsivity. This finding suggests that deficits in executive function lead to more risk-taking through impulsive actions. Stronger verbal skills were also found to predict reduced engagement in risk-taking behaviours. The possibility that inner speech is the mechanism through which verbal ability impacts on risk-taking behaviour was not supported by the findings of this study. However it may be that the methodologies employed were inadequate for detecting inner speech deficits, or relationships between inner speech and risk-taking.

Introduction

'Risk-taking' is a term for behaviours that are associated with some probability of harmful consequences (Boyer, 2006). Engagement in risk-taking behaviours increases during adolescence, and peaks during this developmental period (Steinberg, 2007; Steinberg et al., 2008). The Centre for Disease Control and Prevention (CDC) surveyed the risky behaviours that American high school students engaged in and found a high prevalence of smoking cigarettes (23%) and cannabis (20%), engaging in violent behaviours such as fighting (36%), drinking alcohol (43%), and driving while intoxicated (10%) (Eaton et al., 2006).

Developmental neuroscientists hypothesise that risky behaviour in adolescents results from an interaction between the staggered maturation of two neurobiological systems (Casey, Jones, & Hare, 2008; Steinberg, 2008, 2010; Strang, Chein, & Steinberg, 2013). The first of these systems to mature is a socio-emotional system, comprising limbic and paralimbic brain areas including the ventral striatum and ventromedial prefrontal cortex. The second system incorporates areas of the prefrontal, parietal and anterior cingulate cortices and is thought to underpin many executive functions and increasingly assert regulatory control over risky behaviour as it matures (Casey et al., 2008; Steinberg, 2007, 2008; Strang et al., 2013). The initial maturation of the socio-emotional system during adolescence (prior to the full maturation of the regulatory system) has been postulated to result in an increased propensity to seek rewards and thus explains why adolescents are more likely than children to engage in risk-taking behaviours (Casey et al., 2008; Steinberg, 2008, 2010; Strang et al., 2013). The decline in risk taking behaviour that follows as adolescents reach adulthood (Steinberg, 2007, 2008) is hypothesised to coincide with the maturation of the second regulatory neurobiological system.

However, not all adolescents engage in the same amount of risk-taking behaviour (Romer, 2010). For example, the percentages of individuals who engaged in each of the risky behaviours assessed by the CDC's 2005 Youth Risk Behavior Survey (YRBS) (reported above), indicate that the majority of adolescents did not engage in most of these behaviours (Eaton et al., 2006). Indeed, Romer (2010, p. 264) suggests that "individual differences dominate the emergence of such behavior during adolescence". So what are the factors that may lead individuals to be more or less likely to engage in risk-taking behaviours? One hypothesis that

follows from the developmental neurobiological literature outlined above, is that the executive functions are implicated in the suppression of risk-taking behaviour.

Executive functioning and adolescent risk taking

Executive functions were described by Lezak (1983) as processes that are involved in "how" human behaviour is expressed, and are understood to be necessary for appropriate adult conduct (Lezak, 1983). These functions are the behaviours that humans direct towards themselves in order to modify their behaviour, and therefore allow for the "cross-temporal organisation of behaviour toward the future" (Barkely, 2001, pg 25; Fuster, 1995). These functions are used in self-regulation, allowing individuals to "organise and direct behaviour and cognition" (Jurado & Rosselli, 2007, p. 215) so as to change their future outcomes (Barkley, 2001). Although there is a lack of consensus on the exact definition and nature of the components of the executive system, there is broad agreement that these processes allow people to adapt to a constantly changing environment, through organising thoughts in a goal-directed way and inhibiting inappropriate behaviour (Jurado & Rosselli, 2007). Anderson et al. (2001) conceptualise executive functions as encompassing three integrated processes: 'attentional control' (e.g. selective attention), 'cognitive flexibility' (e.g. working memory) and 'goal setting' (e.g. planning) (Anderson, Anderson, Northam, Jacobs, & Catroppa, 2001).

Certain executive functions in particular have been related to adolescent risk-taking, including response inhibition, planning ability and working memory (Pharo, Sim, Graham, Gross, & Hayne, 2011; Steinberg, 2008). There is some evidence that during adolescence, response inhibition, planning ahead and the weighing of risks and rewards, continue to develop towards adult levels of self-regulation of behaviour (Luna & Sweeney, 2004; Steinberg et al., 2008). For example, a cross-sectional investigation into developmental changes in executive functions in an Australian sample of 11 to 15 year olds, found no sex or age differences in overall planning ability (Anderson et al., 2001). However participants were given two attempts to complete the task, and 11 year olds had a greater number of second attempts than the 15 year old group, indicating there are some subtle, gradual increases in planning ability over adolescence. Anderson et al. (2001) also found age differences (but not sex differences) in working memory, with 15 year olds performing better than younger adolescents. An investigation into the development of response suppression (i.e. inhibition), in a cross-sectional American sample of 8 to 30 year olds found that the ability to suppress a

response was present in children, but efficiency (i.e. being able to use this skill consistently) continued to develop into adolescence (Luna, Garver, Urban, Lazar, & Sweeney, 2004).

Individual differences in these executive functions have indeed been found to predict real-life risk taking. Pharo et al (2011) conducted a study in New Zealand in a community sample of 136, 13 to 17 year old adolescents (69 males, 67 females) and 57, 18 to 22 year olds (27 males, 30 females). They assessed executive functions using a battery of neuropsychological tests that measured processes such as verbal fluency, working memory, selective attention and impulse control, and then combined the results of these tests into one composite score. These researchers found that performance on these tests was predictive of real life risk taking independent of age, sex and personality traits (see below) (Pharo et al., 2011). This finding suggests that individual differences in performance of tests of executive function are related to individual differences in real life adolescent risk-taking. Romer et al (2011) assessed working memory ability at three annual assessments in a longitudinal study of 387 American youths (49% male), who were aged 10 to 12 years at their first assessment. An individual's working memory, as measured at the first assessment, was found to predict their risk-taking behaviour at the third annual assessment, with stronger working memory being associated with reduced risk behaviour (Romer et al., 2011).

Impulsivity and adolescent risk taking

Romer et al's (2011) study also investigated the impact of two forms of impulsivity on risk-taking. They distinguished between self-reported 'sensation seeking', which they conceptualised as "the tendency to pursue exciting and novel experience", and the self-reported "tendency to act without thinking and to experience harmful consequences of impulsive decisions" (Romer et al., 2011, p. 1120). An individual's tendency to 'act without thinking' at their first assessment predicted their engagement in risk-taking behaviour during this first assessment, although it was not associated with later emergence of such behaviours (i.e. at the second and third annual assessments). Given the strength of the association at the first assessment, Romer et al (2011) hypothesised that the effects of 'acting without thinking' on risk-taking were already present, when adolescents were aged between 10 and 12 (Romer et al., 2011). Individual differences in personality traits were also implicated in Pharo and colleagues' (2011) study. They administered the Zuckerman and Kulhman Personality Questionnaire - Short Form (Zuckerman, Kuhlman, Joireman, Teta, & Kraft, 1993) which

assesses 'impulsive sensation-seeking', as well as traits such as 'neuroticism-anxiety' and 'sociability', and derived one composite score from this measure. This 'personality' score was found to predict real world risk taking, once gender, executive function and age were taken into account (Pharo et al., 2011). However it is not clear which of the personality traits combined in the composite score were driving this finding. Stanford et al (1996) have also found that more impulsive individuals engage in greater amounts of risk taking behaviour than less impulsive individuals (Stanford, Greve, Boudreaux, Mathias, & L Brumbelow, 1996).

'Impulsivity' is a heterogeneous construct, which is typically understood as a personality trait and appears to include several different traits in the literature, such as sensation-seeking and lack of premeditation (Dick et al., 2010; Whiteside & Lynam, 2001). Steinberg and colleagues (2008) propose that "impulsivity refers to a lack of self-control or deficiencies in response inhibition; it leads to hasty, unplanned behavior" (Steinberg et al., 2008, p. 1765). Barkley (1997) suggests that inhibitory control provides a delay within which other executive functions can then operate, and that it is a lack of this inhibitory control of inappropriate responses that produces impulsive behaviour, which is suggested to include risk-taking (Barkley, 1997; Romer, 2010). Self-reported impulsivity has indeed been found to be associated with deficits in task based measures of cognitive inhibition (e.g. Stroop) in adults (Enticott, Ogloff, & Bradshaw, 2006).

Investigations into age differences in impulsivity indicate that there is a relatively linear decline in impulsivity from childhood, into adolescence and then adulthood (Galvan, Hare, Voss, Glover, & Casey, 2007; Leshem & Glicksohn, 2007; Steinberg et al., 2008). For example, Steinberg et al (2008) investigated age differences in impulsivity, in an American sample of 935 individuals (49% male, 51% female), aged between 10 and 30 years of age. In this cross-sectional study, participants completed a self-report measure of impulsivity (the Barratt Impulsivity Scale (BIS-11) (Patton, Stanford, & Barratt, 1995)), as well as a behavioural index of impulsivity (conceptualised as hasty performance), which was the amount of time a participant took between presentation and initiation of a computer task. The behavioural measure demonstrated significant declines in impulsivity from age 16 on, in a similar pattern to that seen on the self-report measure, where a linear decline in impulsivity was observed between ages 10 and 30 (Steinberg et al., 2008). The behavioural measure demonstrated a significant effect of gender, with females responding more quickly on their first move, but there were no gender differences on the self-report measure (Steinberg et al., 2008).

Patton and colleagues (1995) investigated the factor structure of the BIS-11 and found three second order factors, termed 'motor impulsiveness' (relating to acting on the spur of the moment), 'non-planning impulsiveness' (relating to planning and thinking carefully) and 'attentional impulsiveness' (relating to focusing on the task at hand). Although they expected to find a 'cognitive impulsiveness' factor, they instead found that cognitive items loaded on to all of the factors. This led them to hypothesise that "'thought processes' in general underlie the personality trait of impulsiveness" (Patton et al., 1995, p. 773). Taken in combination with Romer and colleague's (2011) conceptualisation of impulsivity as 'acting without thinking', it is possible that the absence of particular thought processes significantly contributes to impulsivity. Such thought processes might also be conceptualised as the phenomenon of inner speech.

Inner Speech

In Vygotsky's (1962) theory of cognitive development, children initially depend on external instruction to guide their behaviour, and then learn to do this independently using an outer monologue (often termed 'private speech' in the literature), which eventually becomes internalised as inner speech (Vygotsky, 1962). This hypothesis parallels that of Luria (1965), who understood that in the development of the executive functions, certain prelinguistic capacities exist, which allow for monitoring, planning and inhibition, but in addition a new form of executive functions begin to emerge once these basic capacities enter into relations with language abilities (Luria, 1965; Alderson-Day & Fernyhough, in press).

Inner speech has been defined as "the subjective experience of language in the absence of articulation" (Alderson-Day & Fernyhough, in press, p. 1). Although inner speech is understood to be a separate process from the executive function system (e.g. at times it is involved solely in speech production), it is also hypothesised that the inner speech system is recruited as a cognitive tool to support complex cognitive processes (Alderson-Day & Fernyhough, in press). This hypothesis suggests that disrupting inner speech should impact on people's performance on tasks of executive function. The majority of the research investigating this hypothesis has focused on cognitive flexibility (e.g. switching tasks) and planning ability (Alderson-Day & Fernyhough, in press).

Investigations into whether inner speech supports a particular task have typically employed dual-task paradigms, where participants engage in a secondary task that requires verbal processing. For example, 'articulatory suppression' requires participants to say simple distractor words or phrases (which are understood to disrupt inner speech) while completing the primary executive function task (Alderson-Day & Fernyhough, in press).

Several studies have used articulatory suppression with adult participants to conclude that inner speech facilitates switching from one task to another (i.e. cognitive flexibility) (Emerson & Miyake, 2003) and aids problem solving (on the Wisconsin Card Sorting Test) (Baldo et al., 2005). Typically developing adults have been found to be detrimentally affected on planning tasks (i.e. the Tower of London) when compared with adults with autism (Williams, Bowler, & Jarrold, 2012). These findings have been mirrored in studies with youths, as articulatory suppression was found to impact negatively on the performance of typically developing children and adolescents (when compared with adolescents with autism) during a planning task (the Tower of London) (Wallace, Silvers, Martin, & Kenworthy, 2009) and a cognitive flexibility task (using a computer based card sorting task) (Russell-Smith, Comerford, Maybery, & Whitehouse, 2013). Fatzer and Roebers (2012) found that the greater the working memory demands of tasks administered to 6 and 9 year old children, the greater the detrimental impact of articulatory suppression on their performance (Fatzer & Roebers, 2012). This research suggests that inner speech supports certain executive functions.

The use of a verbal secondary task has also been found to result in faster responding on a task measuring behavioural inhibition (Go/No-go) in adult participants (Tullett & Inzlicht, 2010). In addition, participants made more commission errors (pressing a button when they should not have) than omission errors (failing to press a button), during the articulatory suppression condition. Tullett and Inzlicht (2010) concluded that participants' readiness to press the button constituted more impulsive responding. This study suggests that preventing people from using their inner speech reduces their inhibitory control of inappropriate responses, which in turn produces more impulsive behaviour (Barkley, 1997; Steinberg et al., 2008; Tullett & Inzlicht, 2010).

Dual-task, articulatory suppression experiments constitute one of the methods that are increasingly being used to examine individual differences in inner speech (McCarthy-Jones & Fernyhough, 2011; Alderson-Day & Fernyhough, in press). Questionnaires, such as the

Varieties of Inner Speech Questionnaire (VISQ) (McCarthy-Jones & Fernyhough, 2011), introspective techniques such as Descriptive Experience Sampling (DES) (e.g. Hurlburt, Koch, & Heavey, 2002) and other experimental methods investigating covert lip movements through electromyography (EMG) (e.g. Livesay, Liebke, Samaras, & Stanley, 1996) or silent phonological judgements (e.g. Oppenheim & Dell, 2008) can also be used to examine individual differences in inner speech.

The VISQ measures self-reported phenomenological properties of inner-speech (McCarthy-Jones & Fernyhough, 2011). These specifically investigate the dialogical quality of inner-speech (i.e. the sense of an interplay between two internalised perspectives), its condensed versus expanded nature (as inner speech is hypothesised to be syntactically and semantically abbreviated), its evaluative/motivational nature (whether it is used to evaluate the environment, people and oneself) and whether the voices of other people feature in one's inner speech (McCarthy-Jones & Fernyhough, 2011). This questionnaire has been validated in a British adult undergraduate sample, (McCarthy-Jones & Fernyhough, 2011), but has not previously been used in an adolescent population. There is no comparable valid adolescent measure available.

The current study

Whether an individual does or does not engage in risk-taking behaviour is hypothesised to result from a complex process involving the interplay between inner speech, impulsivity and executive functions. The primary aim of this study is to investigate the degree to which individual differences in these processes predict the amount of real-life risk-taking that adolescents engage in. It is hypothesised that the presence of a situation in which risk-taking could occur will activate the executive system, which will in turn recruit the inner speech system to support its functioning. It is therefore hypothesised that individual differences in both executive functioning and inner speech will contribute to an individual's level of impulsivity, increasing or reducing the likelihood of real-life risk-taking behaviour.

In order to investigate the role of inner speech in risk-taking behaviour, this study will use a proxy measure of adolescent real-life risk-taking in a dual-task paradigm. The adolescent Balloon Analogue Risk Task (BART) is a computer task that requires participants to engage in simulated risk-taking behaviour, during which points can be collected by pumping up a series

of simulated balloons (which increase in value with increasing size) and saving them to a bank before they explode (Lejuez, Aklin, Bornovalova, & Moolchan, 2005). The BART has been found to predict self-reported engagement in real-world risk-taking behaviours in adolescents (Aklin, Lejuez, Zvolensky, Kahler, & Gwadz, 2005; Lejuez et al., 2005; Lejuez, Aklin, Jones, et al., 2003; Lejuez, Aklin, Zvolensky, & Pedulla, 2003). Participants will complete the BART under articulatory suppression and a control condition that involves mouth movements but no speech. This condition is included in order to control for general effects of completing a dual-task (e.g. keeping time to a beat), and identify specific effects of blocking inner speech on risk-taking propensity (Alderson-Day & Fernyhough, in press). It is hypothesised that the greater the difference between a participant's mouth movement and articulatory suppression scores, the more inner speech they will have engaged in during the control task condition.

The VISQ (McCarthy-Jones & Fernyhough, 2011) will also be used to provide a self-report measure of inner speech, in order to both explore and complement findings from the objective articulatory suppression paradigm. Relationships between the subscales on the VISQ and inner speech as measured by the dual-task paradigm are exploratory, as are relationships between the VISQ subscales and the propensity to engage in risk-taking behaviours. Therefore hypotheses about the direction of these relationships will not be made a priori. Participants will also complete a number of measures of executive functioning, and self-reported measures of impulsivity and real-life risk-taking.

Romer et al (2011) found that engaging in risky behaviour increases from the ages of 11 to 14, while other authors report that risk-taking starts to decline in later adolescence, around 18 years of age (Arnett, 1992; Boyer, 2006; Steinberg, 2007, 2008). For this reason, this study will focus on adolescents aged between 13 and 15. In addition, only males will be included as they have reported significantly greater risk-taking behaviours than females (Pharo et al., 2011; Romer et al., 2011).

Hypotheses

1. Performance on the mouth movement BART condition (where inner speech is not disrupted) will be correlated with real-life risk-taking behaviours.
2. Real-life risk taking will be predicted by increased impulsivity, but negatively associated with executive functioning and inner speech (as measured by the

difference in performance between the two computer trials) (please refer to Hypothesis 2 in Figure 1).

3. If a relationship is found between inner speech and real-life risk-taking, this relationship will be mediated by executive function (please refer to Hypothesis 3 in Figure 1).
4. If a relationship is found between executive function and real-life risk-taking this will be mediated by impulsivity (please refer to Hypothesis 4 in Figure 1).

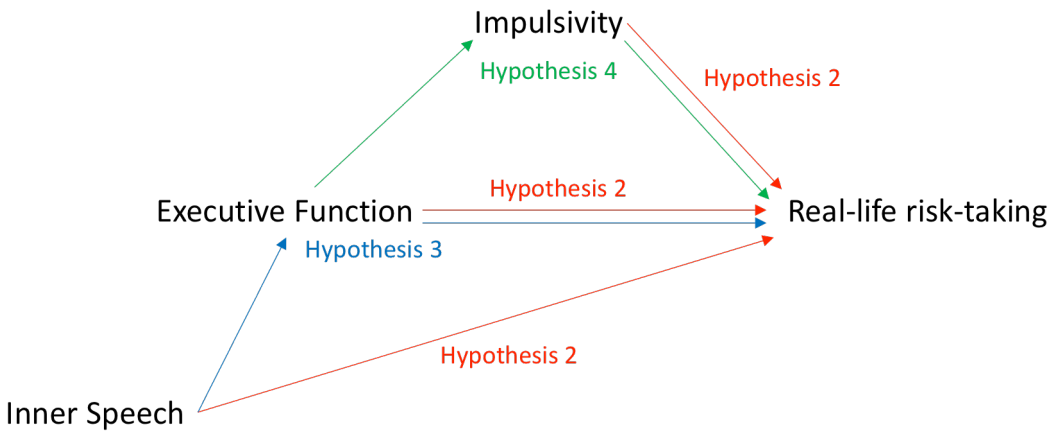


Figure 1_Hypotheses

Method

Participants

A total of sixty, 13-15 year old adolescent males (mean age = 14.74, SD = 0.76) were included in this study.

Inclusion criteria were that participants needed to be male and aged between 13 and 15. Additional inclusion criteria included speaking fluent English and exclusion criteria included having a learning disability or neurological disorder. One participant was excluded because he had a neurological disorder (dyspraxia) and one was excluded because of missing data (not completing questionnaires).

Thirty-six participants were recruited through mainstream schools and twenty-four were recruited from pupil referral units. An opt-in or opt-out parental consent process was employed (see Appendices 1 - 2), depending on the school's discretion. Youth assent was received from all participants (see Appendix 3), in accordance with the protocol approved by the ethics board of King's College London (see Appendix 4). Participants received a £5 Amazon voucher for their time.

Measures

Real-life risk taking

There is no gold-standard measure for assessing real-life risk-taking. Romer et al (2009) assessed real-life risk-taking behaviour using a subset of questions from the USA's Center for Disease Control's (CDC) Youth Risk Behaviour System (YRBS) and the Monitoring the Future Study. MacArthur et al (2012) assessed real-life risk taking behaviour using questions that were initially derived from the Edinburgh Young People's Survey of Crime and then included in the Avon Longitudinal Study of Parents and Children (ALSPAC). These questionnaires ask about behaviour in large epidemiological samples and are therefore not validated per se, although they are well-established in their respective countries.

Relevant questions from both of these studies (MacArthur et al., 2012; Romer et al., 2009) were used to assess real-life risk-taking behaviours. While potentially relevant though ethically challenging, questions relating to sexual behaviours were not included, because of the low prevalence of this behaviour in MacArthur et al's (2012) study (endorsed by 1.3% of their UK sample of 2591 adolescents). The CDC's YRBS questions were used because these behaviours have been assessed by previous studies into adolescent risk-taking (Romer et al., 2009; Romer et al., 2011). However these studies were conducted using American samples, and so questions from the ALSPAC, which also asked about antisocial behaviour, were included as they have been used previously in an epidemiological study in the UK (MacArthur et al., 2012). Selected questions covered the following areas: safety, tobacco, alcohol and drug use, fighting, gang involvement, and antisocial behaviour (see Appendix 7). This measure has clear face validity and is considered to be the best available measure as the questions have been equated to known-risk taking in these previous studies.

Inner speech

Articulatory Suppression and Simulated Risk-Taking with the BART

The adolescent Balloon Analogue Risk Task (BART) is a simulated risk-taking computer task, during which participants pump up a series of balloons (Lejuez et al., 2005). Every pump increases the value and size of the balloon, until the balloon explodes. Participants can choose to collect the points by transferring them into to a permanent bank at any time, as long as the balloon does not explode.

The primary outcome measure on the BART is the adjusted number of pumps across all balloons (Lejuez, Aklin, Zvolensky, et al., 2003). This adjusted value is defined as the average number of pumps excluding balloons that exploded, i.e. the average number of pumps on the balloons before they are transferred to the bank. This value is used as the number of pumps is constrained on balloons that exploded, so including these trials would limit between-participant variability (Lejuez, Aklin, Zvolensky, et al., 2003). This outcome measure has been found to predict self-reported engagement in real-world risk-taking behaviours in adolescents (Aklin et al., 2005; Lejuez et al., 2005; Lejuez, Aklin, Jones, et al., 2003; Lejuez, Aklin, Zvolensky, et al., 2003).

Participants completed the BART under articulatory suppression and a control condition. The articulatory suppression condition followed the procedure used by Wallace et al (2009), with participants instructed to say a word (''Tuesday'') to the beat of a metronome (one beat per second) during task completion. On the rare occasion that a participant stopped saying ''Tuesday'', the experimenter prompted him by saying ''don't forget to say 'Tuesday''' (Wallace et al., 2009).

The articulatory suppression task used resources above and beyond the resources that would be used by inner speech alone, such as keeping to a beat, and moving particular muscles for speech production. Therefore in the control condition, participants were instructed to move their mouths in a sequence that mimicked speech movements (pursing their lips and then smiling) to the beat of the metronome (one beat per second). If a participant stopped moving their mouth the experimenter prompted him by saying "don't forget to move your mouth". Russell-Smith et al. (2013) used a similar procedure so as to measure the performance cost of

mouth movement to a beat and thereby isolate the effect of blocking inner speech on task performance more precisely.

Inner speech while completing the BART computer task was calculated by looking at the adjusted value on the mouth movement (control) trial minus the adjusted value on the articulatory suppression trial. If participants engaged in more pumps (i.e. took more risks) during articulatory suppression than under the control condition, it is suggested that individuals were using their inner speech to control their risk-taking on the BART task. That is, the smaller the value of the difference between these two conditions, the more participants were using their inner speech to control their risk taking (number of pumps) during the control condition.

The Varieties of Inner Speech Questionnaire

The Varieties of Inner Speech Questionnaire (VISQ) has been validated in an adult sample of 235 British university students (77 males, mean age of sample = 20.38, SD = 2.9, range = 18-30). There is no adolescent equivalent of this questionnaire, which specifically measures inner speech (as differentiated from overt private speech). This questionnaire was administered to assess a range of qualitative and functional aspects of inner speech, including measures of evaluative/motivational (e.g. "I talk silently to myself, telling myself not to do things"), dialogic (e.g. "I talk back and forward to myself in my mind about things") and condensed (e.g. "I think to myself in words using full sentences") inner speech and the amount that other voices are present in one's inner speech (e.g. "I hear other people's voices nagging me in my head") (McCarthy-Jones & Fernyhough, 2011) (see Appendix 5).

Impulsivity

Impulsivity was assessed using the Barratt Impulsivity Scale (BIS) (Patton et al., 1995), a widely used self-report measure of impulsive personality traits, which has good construct, convergent and discriminant validity (Patton & Stanford, 1995; Romer et al., 2011). Patton et al's (1995) sample consisted of 412 introductory psychology undergraduates (279 females, 130 males; psychiatric patients (54 females, 110 males) and prisoners (73 males) (ages not given for any group). Undergraduates were found to score significantly lower on the total score on the questionnaire than the psychiatric patient group, and male prison inmates were found to score higher than both undergraduates and the psychiatric patient group (Patton et al., 1995). This questionnaire has also been used widely in adolescent samples (e.g. Steinberg

et al., 2008). Blakemore (personal correspondence) has modified these questions to be appropriate for use with adolescents (Mills, Goddings, Clasen, Giedd, & Blakemore, 2014) (see Appendix 6).

Executive Functions

Working Memory

Working memory was assessed using the backwards digit span task from the Weschler Intelligence Scale for Children IV (WISC-IV) (Wechsler, 2003). Norms for the WISC-IV were derived from a validation sample representative of the UK population of children (aged 6 – 16) (Wechsler, 2004). During this task participants repeat numbers in the reverse order from that presented. The backward span was used as this task requires participants to both store the information temporarily, and manipulate it (Waters & Caplan, 2003).

Planning ability

Planning was assessed using the Tower task from the Delis-Kaplan Executive Function System (DKEFS) battery (Delis, Kaplan, & Kramer, 2001b). Participants are presented with physical wooden disks that can be moved across wooden pegs, with which they are required to plan a sequence of moves and create a visually presented pattern (Holland & Low, 2010). The DKEFS was standardised on a nationally representative American sample (aged 8-89), and a subset of this sample also performed the Tower task again. Test-retest analyses indicated that average performance improved from first to second testing with test-retest correlations in the moderate range (r_{12} = .51) (Delis et al., 2001b). Successful performance on the Tower task requires higher-level executive functions (e.g. rule learning), which include spatial planning abilities, and the total achievement score provides a measure of overall performance (Delis, Kaplan, & Kramer, 2001a).

Cognitive inhibition

Cognitive inhibition was assessed using the Colour-Word Interference Test from the DKEFS battery (Delis et al., 2001b) (see above for information regarding the DKEFS standardisation). This test assesses resistance to interference and inhibitory capacity (Archibald & Kerns, 1999). In the cognitive inhibition condition participants are required to name the colour of the ink a

word is printed in as quickly as possible. The word is the name of a colour, which is always different from the ink that it is written in. The time taken reflects a participant's ability to inhibit the more salient automatic task of reading words in order to name the dissonant ink colours effectively (Delis et al., 2001a). A sample aged between 8 and 19 years completed a re-test assessment on this task; the average time between these administrations was 25 (+/- 12.8 days). Test-retest analyses of number of seconds it took to complete the inhibition task were in the high range (r_{12} (test-retest reliability) = .90) (Delis et al., 2001b).

Cognitive Ability

Cognitive ability was assessed using the Wechsler Abbreviated Scale of Intelligence 2nd Edition (WASI-II) (Wechsler, 2011) to provide an estimate of general intellectual ability. The WASI-II was standardised in a nationally representative sample of the United States English-speaking population of individuals aged 6 to 90 years, and norms were derived from this study (Wechsler, 2011).

The WASI-II is comprised of four sub-tests, two of which measure verbal comprehension abilities (vocabulary and similarities) and two that are considered to assess perceptual reasoning abilities (block design and matrix reasoning). It is possible to compute an individual's full-scale IQ from all four subscales, as well as their verbal comprehension abilities (a measure of crystallised intelligence) and their perceptual reasoning abilities (a measure of fluid intelligence) (Wechsler, 2011). Fluid intelligence is considered to refer to people's abilities to solve problems non-verbally and to perceive novel relationships and is thought to be independent of previous learning and experiences, while crystallised intelligence refers to language comprehension and reasoning abilities, which are considered to be more dependent on prior learning and experiences (Horn & Cattell, 1967).

Procedure

Prior to testing participants were provided with verbal and written explanations (see Appendix 3) of the study, and they were assured of their anonymity. They completed a one and a half hour individual assessment consisting of the WASI-II, followed by one of the two trials of the BART under either AS or control conditions (counterbalanced across participants). They then completed the Towers task, the Colour-Word Interference Task, the digit span, and

the second trial of the BART, followed by the three questionnaires (the real-life risk-taking measure, the adapted BIS-11 and the VISQ).

Analysis

Power analysis

Previous studies have found correlation coefficients between impulsivity and risk taking behaviour of 0.66 (Romer et al., 2009) and between a personality measure, which included assessment of impulsivity, and risk taking behaviour of 0.63 (Pharo et al., 2011). A correlation coefficient of -0.33 has previously been observed between a composite value derived from performance on neuropsychological tests of executive function and risk taking behaviour (Pharo et al., 2011). However the executive function battery used in Pharo et al's (2011) study did not measure all executive functions that have been hypothesised to be involved in risk-taking (Steinberg, 2008), such as ability to plan ahead. The current study aimed to measure executive functions that are more relevant to engagement in risk-taking behaviour and so a correlation estimate of 0.35 was used. A power analysis was conducted using GPower, where significance level was equal to 0.05, power equalled 0.8, and the correlation coefficient (effect size) was equal to 0.35. This resulted in a sample size of 61. A correlation coefficient (effect size) of 0.353 was the result of a sample size of 60, with significance level at 0.05 and power at 0.8.

Planned analyses

Correlation and predictive analyses

Correlation and predictive analyses were conducted using SPSS.

Secondary analyses

Two simple mediation analyses were conducted, and follow-up mediation analyses including covariates were also run. Mediation occurs when the relationship between an independent variable and a dependent variable can be explained by their relationship to a third variable. Mediation analyses distinguish between the direct effect, which is the effect of the independent variable on the dependent variable while controlling for the mediator variable, and the indirect effect, which is the effect of the independent variable through the mediating

variable on the dependent variable (Field, 2013). Mediation analyses were conducted using the mediation module, Process (Hayes, 2013) in SPSS. This macro uses bootstrap samples to estimate the significance of the mediator with a 95% bias corrected adjusted (BCa) confidence interval (CI). Hayes (2013) says that if the 95% CI does not include zero then the indirect effect is significant at $p<0.05$.

Results

Calculation of real-life risk-taking score

Questions on the real-life risk-taking questionnaire were coded with regard to risk being present or absent, and positive responses were summed to provide a total real-life risk-taking score. Specifically each question was coded into a binary variable, where 1 indicated that risk was present, and 0 indicated there was no risk (see Appendix 7). This method of summing affirmative responses has been used in previous investigations into adolescent real-life risk taking behaviours (e.g. Lejuez, Aklin, Zvolensky, et al., 2003).

Fighting was coded as 1 if the participant reported engaging in the behaviour in the past 12 months, and as 0 if he had not, following Romer et al.'s (2009) procedure. Information about whether adolescents did or did not engage in alcohol, tobacco or drug use was included. Drinking alcohol was coded as 0 if a participant said he had never drunk more than a few sips of alcohol, and coded as 1 otherwise. Further information on age of onset and frequency of alcohol, tobacco and drug use was not included, in order to ensure that this behaviour was only represented once in the total scores. It was decided not to include information on whether adolescents had been in a car with a driver who had been drinking or whether a participant's family members or friends were in a gang, as the risk behaviour in these questions was being engaged in by someone else, and participants may not have knowingly engaged in the risk behaviour (only later realising that the driver had been drinking for example). That is, the participant's risk behaviour was association alone and it is not clear how much control participants had over this association. Information about whether a participant's gang hurt people was also not included, as it was thought that this was not asking whether the participants themselves were hurting others directly. Finally information about whether a participant made others do things (question 39E) was not included as it was

not clear which behaviours exactly this was referring to. See Appendix 8 for the proportion of the sample who engaged in each of the behaviours included in the total risk-taking score.

Calculation of executive function and inner speech scores

Executive function was calculated by converting participants' scaled scores on the tower task, the cognitive inhibition task and the working memory task into z scores in order to combine them, and then taking the average of these three scores. Inner speech was calculated by taking away a participant's average adjusted pumps score on the articulatory suppression condition from their average adjusted pumps score on the mouth movement condition.

Order of completing the computer task did not have any effect on performance on the articulatory suppression manipulation, with the difference between participants who completed articulatory suppression first (M=31.52, SE=2.31) and those who completed it second (M=35.8, SE=3.15) not reaching significance, $t(58) = -1.01$, $p=0.32$. There were no significant order effects for the mouth movement condition either, with the difference between those participants who completed the condition first (M=32.37, SE=2.23) and those who completed it second (M=33.5, SE=2.3) also not reaching significance, $t(58) = -0.36$, $p=0.72$.

Descriptive statistics

Table 1 reports descriptive statistics for the risk-taking, executive function composite and individual tasks, impulsivity, inner speech and full scale and composite IQ variables. Participants' total binary real-life risk-taking score was negatively skewed in the sample, and so bias corrected accelerated (BCa) bootstrapping at 1000 samples was used for all regression analyses using this variable.

This total real-life risk-taking score was significantly positively correlated with age, $r = 0.30$, $p < 0.05$. Impulsivity did not correlate with age, $r=0.12$, $p =0.91$, and neither did executive function, $r=-0.04$, $p = 0.78$, inner speech, $r = -0.15$, $p =0.27$, or full scale IQ, $r = -0.12$, $p = 0.34$, verbal comprehension, $r = -0.11$, $p = 0.4$, and perceptual reasoning, $r= -0.12$, $p = 0.37$.

Table 1. Descriptive statistics

Variable	*M*	*SD*	*Range*
Total risk-taking	8.4	5.2	1-22
Executive function	0	0.69	-1.58 – 1.67
Planning	9.9	2.18	5 - 16
Cognitive Inhibition	9.37	3.01	1 – 15
Working Memory	9.88	3.15	3 - 18
Impulsivity	58.92	9.74	38 - 79
Inner Speech	-0.96	0.66	-35.5 – 27.3
Full Scale IQ	99.93	13.9	74 - 141
Verbal Comprehension	100	12.87	82 - 147
Perceptual Reasoning	99.25	14.59	60 - 133

There was a significant effect of source of recruitment on all variables measured, except for inner speech. Specifically, when compared to participants who were collected through comprehensive schools, participants from pupil referral units reported more real-life risk-taking, $t(58) = 3.38$, $p < 0.001$, and impulsivity, $t(58) = 2.12$, $p < 0.05$, and performed less well on measures of executive function, $t(58) = -3.85$, $p < 0.001$, and IQ, $t(58) = -5.03$, $p < 0.001$, as well as on the IQ composites of VCI, $t(58) = -4.05$, $p < 0.001$ and PRI, $t(58) = -4.61$, $p < 0.001$. However this study was interested in within participant differences, rather than between participant differences, and so further analyses did not include source of recruitment.

For each of the questionnaires used in this study, internal reliabilities of the included items were calculated. Specifically, there were 31 items used to calculate participants' real-life risk-taking score ($\alpha = .85$), while the Barrat Impulsivity scale consisted of 27 items ($\alpha = .51$). The VISQ consisted of 18 items ($\alpha = .83$), while the condensed inner speech subscale of the VISQ consisted of 5 items ($\alpha = .17$), the dialogic inner speech subscale consisted of 4 items ($\alpha = .82$), the other people in inner speech subscale consisted of 5 items ($\alpha = .85$) and the evaluative motivational subscale consisted of 4 items ($\alpha = .72$).

Primary Analyses

Correlational analyses

Hypothesis 1 - Performance on the BART (mouth movement condition) will be correlated with real-life risk-taking behaviours.

Performance on the BART computer task did not correlate with real-life risk-taking. This correlation was not significant when using the adjusted average pumps (i.e. average pumps on non-exploded balloons) from either the mouth movement condition, $r = -0.04$, $p = 0.77$, or the articulatory suppression condition, $r = 0.04$, $p = 0.79$.

Exploratory VISQ analyses

The size of participants' inner speech scores did not correlate with any of the VISQ subscales, including the condensed, $r = -0.15$, $p = 0.25$, dialogic, $r = -0.08$, $p = 0.55$, other people, $r = -0.22$, $p = 0.09$ or evaluative/motivational subscales, $r=-0.04$, $p=0.77$. The VISQ subscales also did not correlate with real-life risk-taking behaviour, including the condensed, $r = 0.2$, $p = 0.13$, dialogic, $r = -0.12$, $p = 0.37$, other people, $r = -0.08$, $p = 0.57$ or evaluative/motivational subscales, $r=-0.02$, $p=0.85$.

Predictive analyses

Hypothesis 2 – Real-life risk taking will be predicted by increased impulsivity, but negatively associated with better executive functioning and inner speech.

Three simple regression analyses were run in order to test whether executive function, impulsivity and inner speech each predicted real-life risk-taking. Executive function was found to significantly predict risk-taking behaviour, $F(1, 58) = 6.83$, $p < 0.05$, and this relationship was negative with risk-taking behaviour increasing as executive function decreased (see Appendix 9 for an analysis of the relationship between the components of the executive function composite score and risk-taking). Impulsivity also predicted risk-taking, $F(1, 58) = 21.44$, $p < 0.0001$, with increasing impulsivity leading to increased risk behaviours. These relationships were in the predicted directions.

However inner speech did not predict real-life risk, $F(1,58) = 0.56$, $p = 0.46$. Table 2 shows the results of these simple regression analyses. In addition, there was no significant difference

between the numbers of pumps (i.e. the amount of risk) that participants' engaged in on the mouth movement condition (M=32.92, SD=12.31), when compared to the articulatory suppression condition (M=33.87, SD=15.31), t(59) = 0.71, p = 0.48.

Table 2. Regression analysis for variables predicting real life risk taking.

Predictor Variable	*B*	*95% BCa CI*	*SE B*	*Beta*	*p*
EF prediction of real-life risk-taking					
Constant	8.42	7.05, 9.84	0.67		0.001
Executive Function	-2.43	-4.23, -0.23	0.96	-0.33	0.01
Impulsivity prediction of real-life risk-taking					
Constant	-7.80	-14.60, -0.94	3.35		0.025
Impulsivity	0.28	0.16, 0.39	0.06	0.52	0.001
Inner speech prediction of real-life risk-taking					
Constant	8.37	7.05, 9.8	0.68		0.001
Inner speech	-0.05	-0.19, 0.11	0.08	-0.1	0.46

Simple linear models of predictors of real-life risk taking. Confidence intervals and standard errors based on bias corrected accelerated 1000 bootstrap samples.

Secondary Analyses

Hypothesis 3 - If a relationship is found between inner speech and real-life risk-taking, this relationship will be mediated by executive function.

Although no relationship was found between inner speech and real-life risk-taking, a mediation analysis was still run to investigate whether there was any indirect effect of executive function on the relationship between inner speech (the independent variable) and real-life risk taking (the dependent variable). However this analysis indicated that there was no significant effect of inner speech on risk-taking behaviours through executive function, b = -0.02, BCa CI (-0.11, 0.02).

Hypothesis 4 - If a relationship is found between executive function and real-life risk-taking this will be mediated by impulsivity.

A second mediation analysis was run in order to investigate whether impulsivity was a mediator of the significant relationship between executive function (the independent variable) and real-life risk-taking (the dependent variable). There was a significant indirect effect of executive function on real-life risk-taking through impulsivity, b = -1.26, bias corrected adjustment BCa CI (-2.14, -0.41). This represents a relatively small effect, kappa squared = 0.16, BCa CI (0.05, 0.3) (see Figure 1 for regression coefficients).

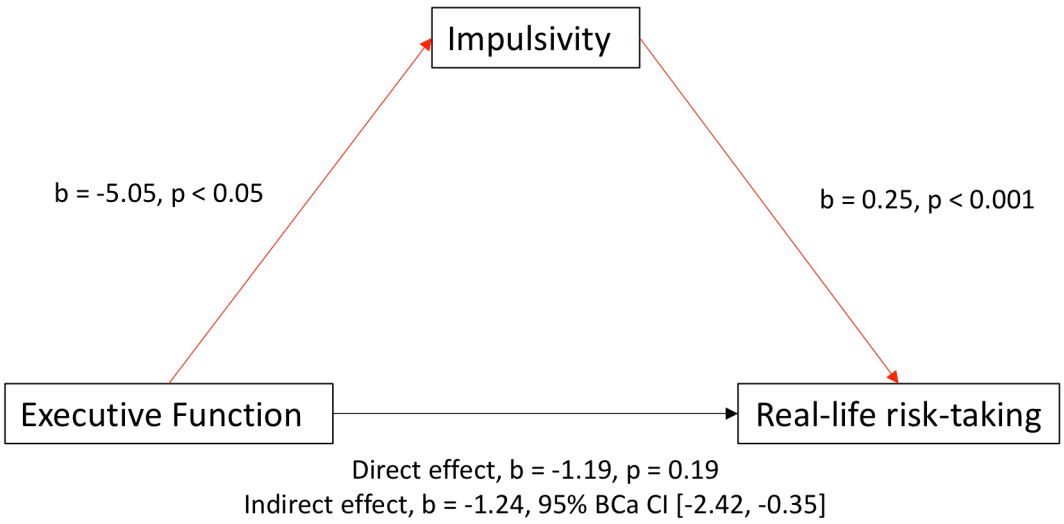


Figure 2_Model of executive function as predictor of risk-taking, mediated by impulsivity. Confidence interval for indirect effect is a BCa bootstrapped CI based on 1000 samples.

Further analyses including IQ

IQ as covariate in mediation analysis

In order to investigate whether IQ had an effect on risk-taking in the above mediation model, a further mediation analysis was run with IQ entered as a covariate. When IQ was controlled for, the significant indirect effect of impulsivity on the relationship between executive function and real-life risk-taking remained, b = -0.70, BCa CI (-1.75, -0.01) (see Figure 2 for regression coefficients).

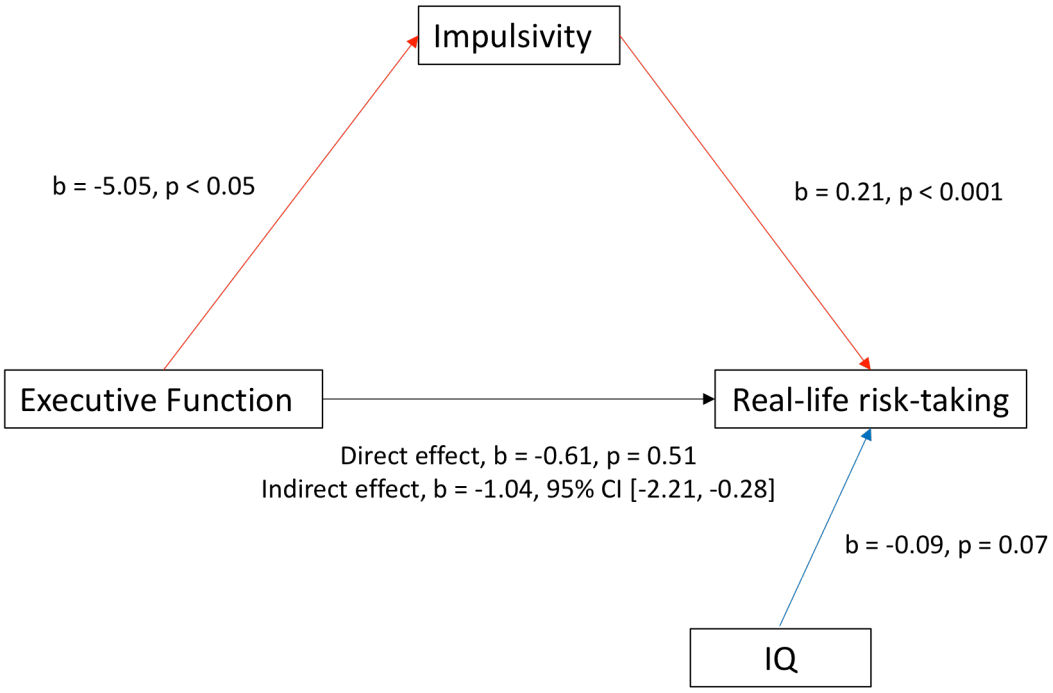


Figure 3_Model of executive function as predictor of risk-taking, mediated by impulsivity, while controlling for IQ. Confidence interval for indirect effect is a BCa bootstrapped CI based on 1000 samples.

A follow-up analysis was then run to investigate whether full-scale IQ predicted risk-taking behaviours. This model was significant, $F(1,58) = 14.89$, $p < 0.001$, and this relationship was negative, with risk-taking behaviour increasing as IQ decreased. Two further simple regression analyses were run to investigate whether the composite IQ scores predicted risk-taking. Both models were also significant, with increasing verbal comprehension predicting less risk-taking, $F(1,58) = 13.17$, $p < 0.001$ and increasing perceptual reasoning also predicting less risky behaviour, $F(1,58) = 9.93$, $p < 0.001$. See Table 3 for the results of these simple regression analyses.

Table 3. Regression analysis of IQ predicting real life risk taking					
Predictor Variable	*B*	*95% BCa CI*	*SE B*	*Beta*	*p*
Full Scale IQ prediction of real-life risk-taking					
Constant	25.18	16.22, 34.58	4.10		0.001
Full Scale IQ	-0.17	-0.24, -0.10	0.04	-0.45	0.001
VCI prediction of real-life risk-taking					
Constant	25.80	17.48, 36.07	4.46		0.001
VCI	-0.17	-0.26, -0.10	0.04	-0.43	0.001
PRI prediction of real-life risk-taking					
Constant	21.84	13.25, 30.96	4.53		0.001
PRI	-0.14	-0.22, -0.05	0.04	-0.38	0.003

Simple linear models of predictors of real-life risk taking. Confidence intervals and standard errors based on bias corrected accelerated 1000 bootstrap samples.

Given that both of the IQ composite scores significantly predicted risk-taking behaviours, a multiple regression analysis was run to investigate whether there was a difference in the predictive strength of an individual's verbal comprehension, compared with their perceptual reasoning abilities, on their real-life risk-taking. These composites were entered into the regression model using the forced entry method, where both predictors are forced into the model simultaneously. In predicting real-world risk-taking, the model was significant overall $F(2,57) = 7.27$, $p < 0.01$. The verbal comprehension subscale was found to predict risk-taking behaviour after controlling for the perceptual reasoning subscale. Table 4 shows the results of this regression analysis.

Table 4. Forced entry regression of IQ composite scores predicting real life risk taking					
Predictor Variable	*B*	*95% BCa CI*	*SE B*	*Beta*	*p*
Constant	27.37	18.16, 37.34	4.52		0.001
VCI	-0.13	-0.23, -0.03	0.05	-0.32	0.019
PRI	-0.06	-0.16, 0.03	0.06	-0.18	0.25

Confidence intervals and standard errors based on bias corrected accelerated 1000 bootstrap samples.

In conclusion, although full-scale IQ predicted risk-taking behaviour, there were differential effects of the composite IQ scores, and PRI did not uniquely predict real-life risk-taking once VCI had been taken into account.

Discussion

Neither performance on the mouth movement nor the articulatory suppression conditions of the BART computer task were associated with real-life risk-taking behaviour. The difference between the mouth movement and the articulatory suppression conditions was hypothesised to be smaller for people who used more inner speech to control their risk-taking on the BART. However the difference between these two conditions was not significant, indicating that there was no effect of inner speech on people's engagement in simulated risk-taking. The difference between the control condition and articulatory suppression conditions (which was hypothesised to relate to the amount of inner speech individuals' engaged in with decreasing difference values relating to increased use of inner speech) was not correlated with any of the subscales of the VISQ. In addition none of the VISQ subscales were correlated with real-life risk-taking behaviours.

Executive function and impulsivity were found to predict engagement in real-life risk-taking behaviours, although inner speech on the BART was unrelated to risk-taking. There was an indirect effect of impulsivity on the relationship between executive functioning and risk-taking behaviour. The indirect effect of executive function on real-life risk-taking through impulsivity remained significant when controlling for IQ. Full scale IQ was also found to predict risk-taking. This relationship was driven by the verbal comprehension composite of the full-scale IQ, as perceptual reasoning did not uniquely predict risk-taking once verbal comprehension was taken into account.

Executive function, impulsivity, IQ and risk-taking.

This study found that poorer executive function abilities, specifically relating to cognitive inhibition, planning ability and working memory, were linked to increased risk-taking. This is consistent with the findings of Pharo and colleagues (2011), who found that individual differences in a composite score of a battery of executive function measures (including working memory, cognitive inhibition and cognitive flexibility) predicted risk-taking behaviour.

Romer and colleagues (2011) also found that working memory performance was inversely related to subsequent risk-taking. Specifically, they found that if an individual had stronger working memory as measured during their first assessment (when aged 10 to 12), this predicted that they would have lowered levels of risk behaviour when it was measured 3 years later (Romer et al., 2011). These relationships support the hypothesis that individual differences in executive functions contribute to their ability to apply inhibitory control to their behaviour, such that poorer executive abilities lead to increased engagement in risk-taking (Romer et al., 2011; Steinberg, 2008).

The finding that self-reported impulsivity predicts risk behaviours is also consistent with previous literature. Romer et al (2009) found that impulsivity, or 'acting without thinking' as they termed it, was strongly related to initiation of risk-taking in the case of young people between the ages of 10 and 12 (Romer et al., 2009), and that it continued to be highly related to risk-taking behaviours at annual assessments of these same adolescents (Romer et al., 2011). It appears that individual differences in adolescent males' levels of impulsivity forms a major contributor to their engagement in risky behaviours.

The current study also found a significant indirect effect of executive function on real-life risk-taking through impulsivity. Impulsivity is considered by some theorists to be related to deficits in inhibitory control within the executive function system (Barkley, 1997), so that people who have high levels of impulsivity have difficulty in controlling inappapropriate actions such as impulsive behaviours (Romer, 2010). The findings in the current study support this hypothesis, given that cognitive inhibition was one of the abilities included in the executive function composite, and suggest that poorer inhibitory control leads to increased risk-taking through impulsive actions. However, given the strength of the mediation effect, which resulted in the direct effect of executive function on risk-taking no longer reaching significance, it may be that the measure of impulsivity used here better captures the construct that is predicting risk-taking, than the measure of executive function that was used.

This indirect effect of executive function on real-life risk-taking through impulsivity remained, once the effect of IQ on risk-taking was controlled for, indicating that this effect was not being driven by intellectual abilities alone. However, this study did find that verbal comprehension abilities inversely predicted engagement in risk-taking. Language ability, including vocabulary, has previously been associated with attentional and behavioural

regulation in adolescents (Petersen et al., 2013; Rodriguez, Mischel, & Shoda, 1989). For example, Peterson et al (2013) analysed externalising behaviours in two separate longitudinal datasets of 585 children from ages 7 to 13 (304 males, 281 females), and of 11,506 children from the ages of 4 to 12 years old (5869 males, 5613 females). Externalising behaviours in the first dataset included aggressive and delinquent behaviours (e.g. "steals outside home", "threatens"), and in the second study included antisocial conduct and oppositional problems (e.g. "cheats or lies") (Petersen et al., 2013). They found that language ability had an independent effect on the development of externalising and inattentive-hyperactive difficulties in both samples. Peterson and colleagues (2013) therefore investigated the impact of language ability on engagement in some of the behaviours (e.g. antisocial behaviour) that were also investigated in the current study. The direction of the effect found in Peterson et al.'s (2013) study was stronger from language ability to behavioural difficulties than in the other direction, and the authors concluded this was supportive of "the causal hypothesis that language ability influences attentional and behavioural regulation" (Petersen et al., 2013, p. 553).

The BART and inner speech

The lack of a relationship between performance on the BART and engagement in real-life risk-taking behaviour is surprising given previous findings. The validiation studies into the use of the BART for predicting such behaviours in adolescents also asked about engagement in risk-behaviours that were included in the USA's Center for Disease Control Youth Risk Behavior Surveillance System ((CDC), 2001), including tobacco, alcohol and drug use, safety, carrying a weapon, stealing and being in a physical fight (Lejuez et al., 2007; Lejuez, Aklin, Zvolensky, et al., 2003). However there were some differences between behaviours that Lejuez et al.'s (2003, 2007) studies assessed including gambling and sexual intercourse without using a condom, and the additional behaviours measured in this study from the ALSPAC study (MacArthur et al., 2012), which related more to antisocial behaviours. Lejuez et al.'s (2002) validation study of the BART as a predictor of adult risk-taking also assessed gambling behaviours as one of the risk-taking behaviours (Lejuez et al., 2002). It may be that the BART only predicts some risk-taking behaviours, such as gambling, and is less valid for predicting more antisocial behaviours. Romer et al. (2009) also did not find a relationship between the BART and risk behaviours, although they used a younger sample (aged 10 to 12 years) than the BART was initially validated in. It may alternatively have been that the mouth movement

and articulatory suppression manipulations that were introduced in this study invalidated the BART task; a possibility that cannot be substantiated given that this study did not use a baseline condition.

As far as is known, this is the first study to investigate the effect of articulatory suppression on performance on the BART computer task. Previous research has found that such dual-task verbal articulation of simple distracting words detrimentally impacts on performance during problem solving (e.g. Baldo et al., 2005), cognitive shifting (e.g. Emerson & Miyake, 2003) and planning (e.g. Williams et al., 2012). However articulatory suppression did not significantly impact on participant's performance on the BART relative to their peformance on the mouth movement condition. A previous study has found that articulatory suppression (repeating the word "the") had no impact on participant's performance during a computerised gambling task where they were told to try to win as much money as possible. However performance was negatively impacted during a digit maintenance condition where participants were presented with a string of numbers (e.g. '25341'), and then asked a question about the numbers (e.g. 'which number is next to 2?') (Jameson, Hinson, & Whitney, 2004). Jameson et al (2004) concluded that it was demands placed on the executive system, rather than the verbal system, which were important for disrupting performance on this gambling task. The null finding in the current study does suggest that participants were not using verbal strategies to support their performance on the BART.

In addition, the subscales of the VISQ were not related to real-life risk-taking. This suggests that, whether or not adolescent boys evaluate their behaviour after it has occurred has no impact on whether they do or do not engage in real-life risk-taking. It also suggests that the amount of dialogic inner speech adolescent boys engage in, the presence of other people's voices in their inner speech, and the condensed nature of their inner speech are also unrelated to their risk taking behaviour. These findings should be generalised with caution, given that the VISQ has not yet been validated in this population.

Verbal abilities, behavioural regulation and future research

This study found that stronger verbal skills were predictive of reduced engagement in risk-taking behaviours. It has been hypothesised that it is the use of private or inner speech which could be the mechanism through which an individual's language ability influences their

behaviour (Petersen et al., 2013; Vygotsky, 1962). The current study investigated this hypothesis, by suggesting that such a mechanism worked through interactions between the inner speech system and the executive function system, the latter of which is able to control behaviour.

However this study found that disrupting inner speech processes appeared to have no impact on adolescent males' propensity to take risks on a computer task that simulated risk-taking. It is possible that the dual-task methodology employed to investigate this relationship may have invalidated the BART task itself. Previous studies have used EMG to show that participants exhibit increased covert lip muscle movements while engaging in elicited inner speech (i.e. when they are instructed to say something silently) compared to visualisation or control tasks (Korba, 1990; Livesay et al., 1996). As far as is known, EMG has not yet been used to assess for the presence of spontaneous inner speech. It would be interesting for future research to investigate whether covert lip movements are detectable by EMG during the BART task, when the BART is administered as originally designed. The presence of such movements would suggest that participants are engaging in inner speech during the task.

Given that this study found no relationship between the BART task and real-life risk-taking, the fact that articulatory suppression had no impact on BART performance does not necessarily preclude the possibility that inner speech is involved in real-life risk-taking. It may be that the methodologies employed in this study were inadequate for detecting inner speech deficits, or relationships between inner speech and risk-taking. The possibility that inner speech is the mechanism through which verbal ability affects risk-taking behaviour requires further exploration.

Descriptive Experience Sampling (DES) is a method of introspection that requires individuals to carry a beeper, make notes on their inner experience when the beeper sounds, and then partake in an interview about these moments within 24 hours (e.g. Hurlburt et al., 2002). This raw data is then coded by investigators for the presence of particular characteristics of inner experience, one of which is inner speech. Although this method has not yet been used extensively with youths (see Hurlburt and Schwitzgebel [2007] for an example of using this method with a 9 year old boy), it would be interesting to use DES with an adolescent sample and investigate whether there is a relationship between the frequency with which individuals engage in inner speech and their real-life risk-taking.

DES is a labour intensive method, and requires participants to collect lots of information themselves in their day-to-day lives. This method may be especially difficult to employ with the harder-to-reach population that tends to constitute youths who attend pupil referral units, but who also appear to take the most risks (in this sample at least). A simpler method of investigating the relationship between inner speech variation and risk taking would be to use questionnaires. As yet the VISQ is the only questionnaire which has been developed to measure inner speech specifically (as differentiated from private speech), and this measure does not investigate the self-regulatory functions that inner speech is hypothesised to serve (McCarthy-Jones & Fernyhough, 2011). The Self-Verbalisation Questionnaire (SVQ) (Duncan & Cheyne, 1999) does include a subscale looking at the cognitive, mnemonic and attentional uses of overt self-verbalisations. Although this questionnaire specifically asks about external private speech, it would be interesting to see whether there is a relationship between the extent to which adolescents self-report using private speech to self-regulate, and their engagement in risk-taking behaviour. Significant relationships would suggest there may also be relationships between the use of inner speech to self-regulate and risk taking, given Vygotsky's (1962) theory that private speech is a developmental precursor to inner speech. It should be noted however, that as far as is known, neither the VISQ nor the SVQ have been validated in adolescent populations. Further development of more comprehensive and appropriate questionnaires is clearly needed.

Inner speech is not the only possible way that language ability may be related to behavioural regulation. For example it has been hypothesised that stronger language skills allow children to more effectively communicate their needs to those around them and that behaviour problems may develop when poorer verbal skills interfere with individual's ability to socialise (Keenan & Shaw, 1997; Petersen et al., 2013).

Limitations

This study used a measure of real-life risk-taking that has not been validated or standardised previously, although sections of the questionnaire have been used in other studies (e.g. Romer et al., 2011). Hypotheses of the relationships between executive functioning, impulsivity and risk-taking were derived from relationships that have been found in previous investigations into adolescent risk taking (Pharo et al., 2011; Romer et al., 2009; Romer et al.,

2011). The confirmation of these hypotheses in the current study, suggests that the measure of real-life risk-taking used was valid. It is suggested that future research could continue to develop this risk-taking measure, in order to determine the factor structure of this questionnaire for example, which was not possible within the constraints imposed by the sample size of the current study.

This study also used the VISQ, which has so far been validated only for adults. In addition the VISQ does not measure some established functions of inner speech, such as it's role in self-regulation. Although other questionnaires (e.g. the Self-Verbalization Questionnaire (Duncan & Cheyne, 1999)) do investigate self-regulation, they focus on private speech (i.e. overt speech directed to oneself) rather than inner speech specifically. It would be useful for future studies to investigate whether the VISQ is valid for use in adolescent populations. It may also be useful to develop a questionnaire that investigates the use of inner speech specifically in self-regulation.

Given the time constraints during testing sessions, a baseline BART condition was not included. Therefore, it was not possible to determine whether the lack of a relationship between the BART and real-life risk-taking behaviours was caused by the introduction of the articulatory suppression and mouth movement conditions. This null finding needs to be validated in future research in this population when using the BART as originally designed. In addition, when using the BART as the experimental task in future articulatory suppression studies, it would be very beneficial to include a baseline performance measure, along with the articulatory suppression and active control conditions, in order to determine whether or not these later conditions invalidate the BART.

This study focused on risk-taking behaviour in adolescent males only, because of the greater prevalence of risk-taking behaviour in this group (e.g. Romer et al., 2011). However, it will be important for future studies to explore whether the findings of the current study hold for female adolescents, and to investigate possible sex differences in larger samples.

Conclusion

This study extends and confirms findings from previous literature on male adolescent risk-taking behaviour (Pharo et al., 2011; Romer et al., 2009; Romer et al., 2011). Stronger

executive function was found to predict reduced engagement in risk-taking behaviour, and this relationship was fully mediated by impulsivity. This finding suggests that deficits in executive function lead to more risk-taking through impulsive actions. As far as is known, this is the first study to demonstrate that stronger verbal skills were predictive of reduced engagement in risk-taking behaviours specifically. This finding adds to a growing literature (e.g. Petersen et al., 2013) suggesting that language abilities are related to behavioural regulation. Inner speech is one mechanism that has been hypothesised to constitute the mechanism by which language skills affect behaviour. Although the current study did not find evidence of inner speech deficits leading to increased risk-taking behaviour, it is possible that this null finding resulted from problems with measurement. Future research should employ different methodologies to investigate the possibility that language skills impact on risk-taking through the mechanism of inner speech.

References

Aklin, W. M., Lejuez, C., Zvolensky, M. J., Kahler, C. W., & Gwadz, M. (2005). Evaluation of behavioral measures of risk taking propensity with inner city adolescents. *Behaviour Research and Therapy, 43*(2), 215-228.

Alderson-Day,B., & Fernyhough,C. (in press). Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology.

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental neuropsychology, 20*(1), 385-406.

Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology, 5*(2), 115-129.

Arnett, J. (1992). Reckless behavior in adolescence: A developmental perspective. *Developmental Review, 12*(4), 339-373.

Baldo, J. V., Dronkers, N. F., Wilkins, D., Ludy, C., Raskin, P., & Kim, J. (2005). Is problem solving dependent on language? *Brain and Language, 92*(3), 240-250.

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychological bulletin, 121*(1), 65.

Barkley, R. A. (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology review, 11*(1), 1-29.

Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review, 26*(3), 291-345.

Casey, B., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences, 1124*(1), 111-126.

CDC (2001). Youth risk behaviour surveillance system. Atlanta, GA: Centers for Disease Control.

Delis, D., Kaplan, E., & Kramer, J. (2001a). D-KEFS: examiners manual. *San Antonio, TX: The Psychological Corporation*.

Delis, D., Kaplan, E., & Kramer, J. (2001b). *Delis-Kaplan executive function system: Technical manual*: Psychological Corporation.

Dick, D. M., Smith, G., Olausson, P., Mitchell, S. H., Leeman, R. F., O'Malley, S. S., & Sher, K. (2010). Review: understanding the construct of impulsivity and its relationship to alcohol use disorders. *Addiction biology, 15*(2), 217-226.

Duncan, R. M., & Cheyne, J. A. (1999). Incidence and functions of self-reported private speech in young adults: A self-verbalization questionnaire. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 31*(2), 133.

Eaton, D. K., Kann, L., Kinchen, S., Ross, J., Hawkins, J., Harris, W. A., . . . Shanklin, S. (2006). Youth risk behavior surveillance—United States, 2005. *Journal of school health, 76*(7), 353-372.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*(1), 148-168.

Enticott, P. G., Ogloff, J. R., & Bradshaw, J. L. (2006). Associations between laboratory measures of executive inhibitory control and self-reported impulsivity. *Personality and Individual Differences, 41*(2), 285-294.

Fatzer, S. T., & Roebers, C. M. (2012). Language and executive functions: The effect of articulatory suppression on executive functioning in children. *Journal of Cognition and Development, 13*(4), 454-472.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*: Sage.

Fuster, J. M. (1995). Memory and planning: Two temporal perspectives of frontal lobe function.

Galvan, A., Hare, T., Voss, H., Glover, G., & Casey, B. (2007). Risk-taking and the adolescent brain: who is at risk? *Developmental science, 10*(2), F8-F14.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*: Guilford Press.

Holland, L., & Low, J. (2010). Do children with autism use inner speech and visuospatial resources for the service of executive control? Evidence from suppression in dual tasks. *British Journal of Developmental Psychology, 28*(2), 369-391.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta psychologica, 26*, 107-129.

Hurlburt, R. T., Koch, M., & Heavey, C. L. (2002). Descriptive experience sampling demonstrates the connection of thinking to externally observable behavior. *Cognitive Therapy and Research, 26*(1), 117-134.

Hurlburt, R. T., & Schwitzgebel, E. (2007). *Describing Inner Experience?: Proponent Meets Skeptic*: Mit Press.

Jameson, T. L., Hinson, J. M., & Whitney, P. (2004). Components of working memory and somatic markers in decision making. *Psychonomic bulletin & review, 11*(3), 515-520.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology review, 17*(3), 213-233.

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological bulletin, 121*(1), 95.

Korba, R. J. (1990). The rate of inner speech. *Perceptual and motor skills, 71*(3), 1043-1052.

Lejuez, C., Aklin, W., Daughters, S., Zvolensky, M., Kahler, C., & Gwadz, M. (2007). Reliability and validity of the youth version of the Balloon Analogue Risk Task (BART–Y) in the assessment of risk-taking behavior among inner-city adolescents. *Journal of Clinical Child and Adolescent Psychology, 36*(1), 106-111.

Lejuez, C., Aklin, W. M., Bornovalova, M. A., & Moolchan, E. T. (2005). Differences in risk-taking propensity across inner-city adolescent ever-and never-smokers. *Nicotine & Tobacco Research, 7*(1), 71-79.

Lejuez, C., Aklin, W. M., Jones, H. A., Richards, J. B., Strong, D. R., Kahler, C. W., & Read, J. P. (2003). The balloon analogue risk task (BART) differentiates smokers and nonsmokers. *Experimental and Clinical Psychopharmacology, 11*(1), 26.

Lejuez, C., Aklin, W. M., Zvolensky, M. J., & Pedulla, C. M. (2003). Evaluation of the Balloon Analogue Risk Task (BART) as a predictor of adolescent real-world risk-taking behaviours. *Journal of adolescence, 26*(4), 475-479.

Lejuez, C., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., . . . Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75.

Leshem, R., & Glicksohn, J. (2007). The construct of impulsivity revisited. *Personality and Individual Differences, 43*(4), 681-691.

Lezak, M. (1983). Neuropsychological assessment: Oxford University Press, New York.

Livesay, J., Liebke, A., Samaras, M., & Stanley, A. (1996). Covert speech behavior during a silent language recitation task. *Perceptual and motor skills, 83*(3f), 1355-1362.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child development, 75*(5), 1357-1372.

Luna, B., & Sweeney, J. A. (2004). The emergence of collaborative brain function: FMRI studies of the development of response inhibition. *Annals of the New York Academy of Sciences, 1021*(1), 296-309.

Luria, A. R. (1965). LS Vygotsky and the problem of localization of functions. *Neuropsychologia, 3*(4), 387-392.

MacArthur, G., Smith, M., Melotti, R., Heron, J., Macleod, J., Hickman, M., . . . Lewis, G. (2012). Patterns of alcohol use and multiple risk behaviour by gender during early and late adolescence: the ALSPAC cohort. *Journal of Public Health, 34*(suppl 1), i20-i30.

McCarthy-Jones, S., & Fernyhough, C. (2011). The varieties of inner speech: links between quality of inner speech and psychopathological variables in a sample of young adults. *Consciousness and cognition, 20*(4), 1586-1593.

Mills, K. L., Goddings, A.-L., Clasen, L. S., Giedd, J. N., & Blakemore, S.-J. (2014). The developmental mismatch in structural brain maturation during adolescence. *Developmental neuroscience, 36*(3–4), 147-160.

Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition, 106*(1), 528-537.

Patton, J. H., & Stanford, M. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology, 51*(6), 768-774.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the Barratt impulsiveness scale. *Journal of clinical psychology*(51), 768-774.

Petersen, I. T., Bates, J. E., D'Onofrio, B. M., Coyne, C. A., Lansford, J. E., Dodge, K. A., . . . Van Hulle, C. A. (2013). Language ability predicts the development of behavior problems in children. *Journal of abnormal psychology, 122*(2), 542.

Pharo, H., Sim, C., Graham, M., Gross, J., & Hayne, H. (2011). Risky Business: Executive Function, Personality, and Reckless Behavior During Adolescence and Emerging Adulthood. *Behavioral Neuroscience, 125*(6).

Rodriguez, M. L., Mischel, W., & Shoda, Y. (1989). Cognitive person variables in the delay of gratification of older children at risk. *Journal of Personality and Social Psychology, 57*(2), 358.

Romer, D. (2010). Adolescent risk taking, impulsivity, and brain development: Implications for prevention. *Developmental Psychobiology, 52*(3), 263-276.

Romer, D., Betancourt, L., Giannetta, J. M., Brodsky, N. L., Farah, M., & Hurt, H. (2009). Executive cognitive functions and impulsivity as correlates of risk taking and problem behavior in preadolescents. *Neuropsychologia, 47*(13), 2916-2926.

Romer, D., Betancourt, L. M., Brodsky, N. L., Giannetta, J. M., Yang, W., & Hurt, H. (2011). Does adolescent risk taking imply weak executive function? A prospective study of relations between working memory performance, impulsivity, and risk taking in early adolescence. *Developmental science, 14*(5), 1119-1133.

Russell-Smith, S. N., Comerford, B. J., Maybery, M. T., & Whitehouse, A. J. (2013). Brief Report: Further Evidence for a Link Between Inner Speech Limitations and Executive Function in High-Functioning Children with Autism Spectrum Disorders. *Journal of autism and developmental disorders*, 1-8.

Stanford, M. S., Greve, K. W., Boudreaux, J. K., Mathias, C. W., & L Brumbelow, J. (1996). Impulsiveness and risk-taking behavior: Comparison of high-school and college students using the Barratt Impulsiveness Scale. *Personality and Individual Differences, 21*(6), 1073-1075.

Steinberg, L. (2007). Risk taking in adolescence new perspectives from brain and behavioral science. *Current Directions in Psychological Science, 16*(2), 55-59.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28*(1), 78-106.

Steinberg, L. (2010). A dual systems model of adolescent risk-taking. *Developmental Psychobiology, 52*(3), 216-224.

Steinberg, L., Albert, D., Cauffman, E., Banich, M., Graham, S., & Woolard, J. (2008). Age Differences in Sensation Seeking and Impulsivity as Indexed by Behavior and Self-Report: Evidence for a Dual Systems Model. *Developmental Psychology, 44*(6), 1764-1778.

Strang, N. M., Chein, J. M., & Steinberg, L. (2013). The value of the dual systems model of adolescent risk-taking. *Frontiers in human neuroscience, 7*, 223-223.

Tullett, A. M., & Inzlicht, M. (2010). The voice of self-control: Blocking the inner voice increases impulsive responding. *Acta psychologica, 135*(2), 252-256.

Vygotsky, L. (1962). *Thought and language*. Cambridge, MA: The MIT Press.

Wallace, G. L., Silvers, J. A., Martin, A., & Kenworthy, L. E. (2009). Brief report: Further evidence for inner speech deficits in autism spectrum disorders. *Journal of autism and developmental disorders, 39*(12), 1735-1739.

Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers, 35*(4), 550-564.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-WISC-IV*: Psychological Corporation.

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children Fourth UK Edition (WISC-IV UK) Administation and Scoring Manual*. London: Harcourt Assessment.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence Second Edition Manual*. Bloomington, MN: Pearson.

Whiteside, S. P., & Lynam, D. R. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences, 30*(4), 669-689.

Williams, D. M., Bowler, D. M., & Jarrold, C. (2012). Inner speech is used to mediate short-term memory, but not planning, among intellectually high-functioning adults with autism spectrum disorder. *Development and Psychopathology, 24*(01), 225-239.

Zuckerman, M., Kuhlman, D. M., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: The Big Three, the Big Five, and the Alternative Five. *Journal of Personality and Social Psychology, 65*(4), 757.

Appendices

Appendix 1: Opt-In Parent Information Sheet and Consent Form

INFORMATION SHEET FOR PARENTS/GUARDIANS

REC Reference Number: PNM/13/14-59

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Individual differences in adolescent risk-taking behaviour: investigating relations between risk-taking, executive function, impulsivity and inner speech.

We would like to invite your son to participate in this postgraduate research project. He should only participate if you want him to; choosing not to take part will not disadvantage him in any way. Before you decide whether you want him to take part, it is important to understand why the research is being done and what his participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. We suggest that you take at least 24 hours to consider whether you would like your child to participate after reading this sheet. Ask us if there is anything that is not clear or if you would like more information.

We know that some teenagers engage in more risk-taking behaviour, like drinking alcohol, than their peers. This project is aiming to find out why some teenagers are more likely to do this than others. The project is being conducted as part of a Doctorate in Clinical Psychology at King's College London (KCL) and it is being funded by the National Health Service (NHS).

To take part in this study, your child needs to be male and between the ages of 13 and 15. He also needs to have a good understanding of written and verbal English, and not have a diagnosis of neurological disease or learning disability. If he does not meet these criteria then his data will not be included in the study, and will be immediately destroyed. It is your decision whether he takes part or not; his taking part in the study is not related to his behaviour. If you and your child decide that he would like to participate then you should sign the below consent form and return it to school. Taking part will teach students about how this kind of research is conducted and it is hoped that they will find it interesting to be a part of this process.

We will arrange a 1 ½ hour appointment with your son at his school, or at the Institute of Psychiatry, KCL. He will be asked to complete a computer task that simulates real-life risk taking and some tasks that look at memory, ability to plan and pay attention. He will also complete a brief assessment of his intellectual ability and two questionnaires. One of these

asks about his behaviour, such as “over the past year have you been rowdy or rude in a public place so that people complained or you got into trouble?”. The other questionnaire asks about your son’s tendency to act impulsively, such as “I do things without thinking”. After the session he will have the opportunity to discuss his results and any thoughts he had during the session. If you or your son would like to receive a copy of the overall study results, this will be arranged.

All the data that your son provides will be treated confidentially, which means it will be private. Only the researchers will have access to it. His data will be anonymised, which means that it will not be stored with any of his personally identifiable information. The only reason we would share information with others is if we felt your son or someone else was at risk (e.g use of weapons). The anonymised findings of this research will form part of a thesis for a doctorate in clinical psychology and will be submitted to a peer-review journal and presented at specialist conferences.

It is up to you and your son to decide whether you would like him to take part in the study or not. If he does decide to take part he will still be free to withdraw from the study without giving a reason and at any time until the thesis has been submitted for examination (i.e. until 01/05/2015). If you do decide to participate then please sign the attached consent form, and ask him to return it to school. This information sheet is yours to keep, and your son will return a copy of the consent form to you.

If you have any questions or require more information about this study, please contact the researcher using the following contact details: Zoe Hyde: zoe.hyde@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF, 0207 848 0223.

NB: If this study has harmed you in any way, you can contact King's College London using the details below for further advice and information: Dr Troy Tranah, troy.tranah@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF.

CONSENT FORM FOR PARTICIPANTS IN RESEARCH STUDIES

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study: Individual differences in adolescent risk-taking behaviour: investigating relations between risk-taking, executive function, impulsivity and inner speech.

King's College Research Ethics Committee Ref: PNM/13/14-59

Thank you for considering taking part in this research. The person organising the research must explain the project to you before you agree for your son to take part. If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher on the contact details on the information sheet before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

Please tick or initial

- I understand that if my son or I decide at any time during the research that he no longer wishes to participate in this project, we can notify the researchers involved and he can withdraw from it immediately without giving any reason. Furthermore, I understand that he is able to withdraw his data up to the point when it is submitted as a part of the researcher's Doctorate in Clinical Psychology thesis. ☐

- I consent to the processing of my son's personal information for the purposes explained to me. I understand that such information will be handled in accordance with the terms of the UK Data Protection Act 1998. ☐

- I understand that confidentiality and anonymity will be maintained and it will not be possible to identify me in any publications. ☐

- Please indicate whether you would like to receive a copy of a report about the findings of the study and provide an email/postal address at which you would like to receive this report: ☐

Parent/Guardian Statement:

I _____

agree that the research project named above has been explained to me to my satisfaction and I agree that my son can take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: _____ Date: _____

Investigator's Statement:

I _____

Confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the participant.

Signed: _____ Date: _____

Appendix 2: Opt-Out Parent Information Sheet and Consent Form

INFORMATION SHEET FOR PARENTS/GUARDIANS

REC Reference Number: PNM/13/14-59

THIS INFORMATION SHEET IS YOURS TO KEEP FOR YOUR RECORDS

Individual differences in adolescent risk-taking behaviour: investigating relations between risk-taking, executive function, impulsivity and inner speech.

We are inviting your son to participate in this postgraduate research project. He will not participate if you do not want him to; choosing not to take part will not disadvantage him in any way. Before you decide whether you do not want him to take part, it is important to understand why the research is being done and what his participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. We suggest that you take the next week to consider whether you do not want your child to participate after reading this sheet. Please contact Zoe Hyde, who is the lead researcher, (see contact details below) if there is anything that is not clear or if you would like more information.

We know that some teenagers engage in more risk-taking behaviour, like drinking alcohol, than their peers. This project is aiming to find out why some teenagers are more likely to do this than others. The project is being conducted as part of a Doctorate in Clinical Psychology at King's College London (KCL) and it is being funded by the National Health Service (NHS).

To take part in this study, your child needs to be male and between the ages of 13 and 15. He also needs to have a good understanding of written and verbal English, and not have a diagnosis of neurological disease or learning disability. If he does not meet these criteria then his data will not be included in the study, and will be immediately destroyed. It is your decision as to whether he does not take part in this study. If you and your child decide that you would not like him to participate then you should sign the opt-out consent form (see below) and return it to school. Taking part will teach students about how this kind of research is conducted and it is hoped that they will find it interesting to be a part of this process. If you are happy for you son to take part then you do not need to do anything.

We will arrange a 1 ½ hour appointment with your son at his school. He will be asked to complete a computer task that simulates real-life risk taking and some tasks that look at memory, ability to plan and pay attention. He will also complete a brief assessment of his intellectual ability and three questionnaires. One of these asks about his behaviour, such as "over the past year have you been rowdy or rude in a public place so that people

complained or you got into trouble?". Another questionnaire asks about your son's tendency to act impulsively, such as "I do things without thinking". The final questionnaire asks questions about the act of talking silently to oneself when thinking, such as "I think to myself in words using full sentences." After the session he will have the opportunity to discuss his results and any thoughts he had during the session. If you or your son would like to receive a copy of the overall study results, this will be arranged (see below).

All the data that your son provides will be treated confidentially, which means it will be private. Only the researchers will have access to it. His data will be anonymised, which means that it will not be stored with any of his personally identifiable information. The only reason we would share information with others is if we felt your son or someone else was at risk (e.g use of weapons). The anonymised findings of this research will form part of a thesis for a doctorate in clinical psychology and will be submitted to a peer-review journal and presented at specialist conferences.

It is up to you and your son to decide whether you would not like him to take part in the study. If he does take part he will still be free to withdraw from the study, without giving a reason, at any time until the thesis has been submitted for examination (i.e. until 01/05/2015).

If you do decide that you do not want your son to participate then please sign the attached opt-out consent form, and ask him to return it to school. This information sheet is yours to keep.

If you have any questions or require more information about this study, please contact the researcher using the following contact details: Zoe Hyde: zoe.hyde@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF, 0207 848 0223.

You can also contact Dr Troy Tranah, who is the Lead Psychologist for CAMHS, and is another researcher on the study. Dr Tranah can be contacted at: troy.tranah@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF.

NB: If this study has harmed you in any way, you can contact King's College London using the following details for further advice and information: Dr Tranah; troy.tranah@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF.

OPT-OUT CONSENT FORM

Please complete this form after you have read the Information Sheet.

Title of Study: Individual differences in adolescent risk-taking behaviour: investigating relations between risk-taking, executive function, impulsivity and inner speech.

King's College Research Ethics Committee Ref: PNM/13/14-59

Thank you for considering taking part in this research. If you have any questions arising from the Information Sheet, please talk to the researcher on the contact details on the information sheet.

If you DO NOT want your son to take part in this research project then please sign this form and ask your son to return it to his school.

Please print your son's name _____

Parent/Guardian Statement:

I _____

do not want my son _____ (print son's name) to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: _____ Date: _____

REQUEST FOR A COPY OF THE OVERALL STUDY RESULTS

If you or your son would like to have a copy of the overall study results then please sign this form and ask your son to return it to his school.

Please tick or initial

- Please indicate whether you would like to receive a copy of a report about the findings of the study and provide an email/postal address at which you would like to receive this report: ☐

_____

_____

Please print your son's name

_____

Appendix 3: Adolescent Information Sheet

INFORMATION SHEET FOR PARTICIPANTS

REC Reference Number: PNM/13/14-59

YOU WILL BE GIVEN A COPY OF THIS INFORMATION SHEET

Individual differences in adolescent risk-taking behaviour: investigating relations between risk-taking, executive function, impulsivity and inner speech.

We would like to invite you to participate in this postgraduate research project. You should only participate if you want to; choosing not to take part will not disadvantage you in any way. Before you decide whether you want to take part, it is important to understand why the research is being done and what your participation will involve. Please take time to read the following information carefully and discuss it with others if you wish. We suggest that you take at least 24 hours to consider whether to participate after reading this sheet. Ask us if there is anything that is not clear or if you would like more information.

We know that some teenagers engage in more risk-taking behaviour, like drinking alcohol, than their peers. This project is aiming to find out why some teenagers are more likely to do this than others. The project is being conducted as part of a Doctorate in Clinical Psychology at King's College London (KCL) and it is being funded by the National Health Service (NHS).

To take part in this study, you need to be male and between the ages of 13 and 15. You need to have a good understanding of written and verbal English, and not have a diagnosis of neurological disease or learning disability. If you do not meet these criteria then your data will not be included in the study, and will be immediately destroyed. It is your decision whether you take part or not; your taking part in the study is not related to your behaviour. Taking part will teach you about how this kind of research is conducted and it is hoped that you will find it interesting to be a part of this process.

We will arrange a 1 ½ hour appointment with you at your school, or at the Institute of Psychiatry, KCL. You will be asked to complete a computer task that simulates real-life risk taking and some tasks that look at memory, ability to plan and pay attention. You will also complete a brief assessment of your intellectual ability and two questionnaires. One of these asks about your behaviour, such as "over the past year have you been rowdy or rude in a public place so that people complained or you got into trouble?". The other questionnaire asks about your tendency to act impulsively, such as "I do things without thinking". After the session you will have the opportunity to discuss your results and any thoughts you had

during the session. If you or your parent/guardian would like to receive a copy of the overall study results, this will be arranged.

All the data that you provide will be treated confidentially, which means it will be private. Only the researchers will have access to it. Your data will be anonymised, which means that it will not be stored with any of your personally identifiable information. The only reason we would share information with others is if we felt you or someone else was at risk (e.g. use of weapons). The anonymised findings of this research will form part of a thesis for a doctorate in clinical psychology and will be submitted to a peer-review journal and presented at specialist conferences.

It is up to you and your parent/guardian to decide whether you would like to take part in the study or not. If you do decide to take part you will still be free to withdraw from the study without giving a reason and at any time until the thesis has been submitted for examination (i.e. until 30/04/2015). If you do decide to participate then please get your parent/guardian to sign the consent form, and return it to school. This information sheet is yours to keep.

If you have any questions or require more information about this study, please contact the researcher using the following contact details: Zoe Hyde: zoe.hyde@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF, 0207 848 0223.

NB: If this study has harmed you in any way, you can contact King's College London using the details below for further advice and information: Dr Troy Tranah, troy.tranah@kcl.ac.uk, Institute of Psychiatry, Addiction Sciences Building, 4 Windsor Walk, London, SE5 8AF.

Appendix 4: Ethical Approval

Zoe Hyde
Addiction Scienes Building
4 Windsor Walk
London
SE5 8AF

25 April 2014

Dear Zoe Hyde

PNM/13/14-59 Individual differences in adolescent risk-tasking behaviour: Investigation relations between risk-taking, executive function, impulsivity and inner speech.

Review Outcome: Full Approval

Thank you for sending in the amendments/clarifications requested to the above project. I am pleased to inform you that these meet the requirements of the PNM RESC and therefore that full approval is now granted.

Please ensure that you follow all relevant guidance as laid out in the King's College London Guidelines on Good Practice in Academic Research (http://www.kcl.ac.uk/college/policyzone/index.php?id=247).

For your information ethical approval is granted until 25/04/2015. If you need approval beyond this point you will need to apply for an extension to approval at least two weeks prior to this explaining why the extension is needed, (please note however that a full re-application will not be necessary unless the protocol has changed). You should also note that if your approval is for one year, you will not be sent a reminder when it is due to lapse.

Ethical approval is required to cover the duration of the research study, up to the conclusion of the research. The conclusion of the research is defined as the final date or event detailed in the study description section of your approved application form (usually the end of data collection when all work with human participants will have been completed), not the completion of data analysis or publication of the results.

For projects that only involve the further analysis of pre-existing data, approval must cover any period during which the researcher will be accessing or evaluating individual sensitive and/or un-anonymised records.

Note that after the point at which ethical approval for your study is no longer required due to the study being complete (as per the above definitions), you will still need to ensure all

research data/records management and storage procedures agreed to as part of your application are adhered to and carried out accordingly.

If you do not start the project within three months of this letter please contact the Research Ethics Office.

Should you wish to make a modification to the project or request an extension to approval you will need approval for this and should follow the guidance relating to modifying approved applications:
http://www.kcl.ac.uk/innovation/research/support/ethics/applications/modifications.aspx

Please would you also note that we may, for the purposes of audit, contact you from time to time to ascertain the status of your research.

If you have any query about any aspect of this ethical approval, please contact your panel/committee administrator in the first instance (http://www.kcl.ac.uk/innovation/research/support/ethics/contact.aspx)
We wish you every success with this work.

Yours sincerely,

Tom Billins, Senior Research Ethics Officer

For and on behalf of
Professor Gareth Barker, Chairman
Psychiatry, Nursing and Midwifery Research Ethics Subcommittee (PNM RESC)

Cc. Troy Tranah

Appendix 5: Varieties of Inner Speech Questionnaire

Author's permission for use in this study only.

Varieties of Inner Speech Questionnaire (McCarthy-Jones & Fernyhough, 2011)

The following questions are about inner speech, or the act of talking silently to oneself when thinking. Please read the statements given below and select the option that most applies to you.

1 = Certainly does not apply to me

2 = Possibly does not apply to me

3 = If anything, slightly does not apply to me

4 = If anything, applies to me slightly

5 = Possibly applies to me

6 = Certainly applies to me

1	I think to myself in words using brief phrases and single words rather than full sentences.	1 2 3 4 5 6
2	When I am talking to myself about things in my mind, it is like I am going back and forward asking myself questions and then answering them	1 2 3 4 5 6
3	I hear the voice of another person in my head. For example, when I have done something foolish I hear my mother's voice criticising me in my mind.	1 2 3 4 5 6
4	I experience the voices of other people asking me questions in my head.	1 2 3 4 5 6
5	I hear other people's voices nagging me in my head.	1 2 3 4 5 6
6	My thinking in words is more like a dialogue with myself, rather than my own thoughts in a monologue.	1 2 3 4 5 6
7	I think to myself in words using full sentences.	1 2 3 4 5 6
8	My thinking to myself in words is like shorthand notes, rather than full, proper, grammatical English.	1 2 3 4 5 6

9	I think in inner speech about what I have done, and whether it was right or not.	1 2 3 4 5 6
10	When I am talking to myself about things in my mind, it is like I am having a conversation with myself.	1 2 3 4 5 6
11	I talk silently to myself telling myself to do things.	1 2 3 4 5 6
12	I hear other people's actual voices in my head, saying things that they have never said to me before.	1 2 3 4 5 6
13	I talk back and forward to myself in my mind about things.	1 2 3 4 5 6
14	My thinking in words is shortened compared to my normal out-loud speech. For example, rather than saying to myself things like 'I need to go to the shops,' I will just say 'shops' to myself in my head.	1 2 3 4 5 6
15	If I were to write down my thoughts on paper, they would read like a normal grammatical sentence.	1 2 3 4 5 6
16	I hear other people's actual voices in my head, saying things that they actually once said to me.	1 2 3 4 5 6
17	I talk silently to myself telling myself not to do things.	1 2 3 4 5 6
18	I evaluate my behaviour using my inner speech. For example I say to myself, 'that was good' or 'that was stupid.'	1 2 3 4 5 6

Appendix 6: Adolescent Barratt Impulsivity Scale

Author's permission for use in this study only.

*People differ in the ways they act and think in different situations. This is a test to measure some of the ways in which **you tend to think and act**. Read each statement and **circle the appropriate number on the right side of this page**. Please answer all the statements and do not leave any of the statements blank.*

		Rarely/ Never	Occa-sionally	Often	Usually/ Always
1.	I plan tasks carefully.	1	2	3	4
2.	I do things without thinking.	1	2	3	4
3.	I make up my mind quickly.	1	2	3	4
4.	I am happy-go-lucky.	1	2	3	4
5.	I don't "pay attention."	1	2	3	4
6.	I have "racing" thoughts.	1	2	3	4
7.	I plan trips well ahead of time.	1	2	3	4
8.	I am self-controlled.	1	2	3	4
9.	I can concentrate easily.	1	2	3	4
10.	I save my money regularly.	1	2	3	4
11.	I can't keep still in school lessons.	1	2	3	4
12.	I am a careful thinker.	1	2	3	4
13.	I plan my career.	1	2	3	4
14.	I say things without thinking.	1	2	3	4
15.	I like to think about complex problems.	1	2	3	4
16.	I change interests or hobbies frequently.	1	2	3	4
17.	I act "on impulse."	1	2	3	4
18.	I get easily bored when solving thought problems.	1	2	3	4
19.	I act on the spur of the moment.	1	2	3	4
20.	I am a steady thinker.	1	2	3	4
21.	I buy things on impulse.	1	2	3	4
22.	I can only think about one thing at a time.	1	2	3	4
23.	I often have unrelated thoughts while thinking.	1	2	3	4
24.	I am more interested in the present than the future.	1	2	3	4
25.	I am restless at the theatre or the cinema .	1	2	3	4
26.	I like puzzles.	1	2	3	4
27.	I am future-orientated.	1	2	3	4

Appendix 7: All risk-behaviours and binary coding of included risk questions.

These questions ask about safety.

1. When you ride a bicycle, how often do you wear a helmet?

A. I do not ride a bicycle	0
B. Never wear a helmet	1
C. Rarely wear a helmet	1
D. Sometimes wear a helmet	1
E. Most of the time wear a helmet	1
F. Always wear a helmet	0

2. When you rollerblade or ride a skateboard, how often do you wear a helmet?

A. I do not rollerblade or ride a skateboard	0
B. Never wear a helmet	1
C. Rarely wear a helmet	1
D. Sometimes wear a helmet	1
E. Most of the time I wear a helmet	1
F. Always wear a helmet	0

3. How often do you wear a seat belt when riding in a car?

A. Never	1
B. Rarely	1
C. Sometimes	1
D. Most of the time	1
E. Always	0

4. Have you ever ridden in a car driven by someone who had been drinking alcohol?

A. Yes	
B. No	
C. Not sure	

These questions ask about violence-related behaviors.

5. Have you ever carried a weapon, such as a gun, knife, or club?

A. Yes	1
B. No	0

6. Have you ever been in a physical fight?

A. Yes	
B. No	

7. Have you ever been in a physical fight in which you were hurt and had to be treated by a doctor or nurse?

A. Yes	1
B. No	0

8. During the past 12 months, how many times were you in a physical fight?

A. 0 times	0
B. 1 time	1
C. 2 or 3 times	1
D. 4 or 5 times	1
E. 6 or 7 times	1
F. 8 or 9 times	1
G. 10 or 11 times	1
H. 12 or more times	1

The next questions ask about drinking alcohol.

18. Have you ever had an alcoholic drink?

A. Yes	
B. No	

19. How old were you when you had you first drink of alcohol (other than a few sips)?

A. I have never had a drink of alcohol other than a few sips	0
B. 10 years old or younger	1
C. 11 years old	1
D. 12 years old	1
E. 13 years old	1
F. 14 years old	1
G.15 years old	1

20. During the last 12 months how many days did you drink alcohol?

A.1 day	
B. 2-5 days	
C. 6-9 days	
D. 10-19 days	
E. 20-30 days	
F. 30-40 days	
G. 40 or more days	

21. Have you ever been sick or felt unwell from drinking alcohol?

A. Yes	1
B. No	0

The next questions ask about smoking tobacco and cigarettes.

22. Have you ever tried smoking, even on or two puffs?

A. Yes	1
B. No	0

23. How old were you when you smoked a whole cigarette for the first time?

A. I have never smoked a cigarette	
B. 10 years old or younger	
C. 11 years old	
D. 12 years old	
E. 13 years old	
F. 14 years old	
G. 15 years old	

24. During the past 12 months, on how many days did you smoke a cigarette?

A. 0 days	
B. 1-5 days	
C. 6-20 days	
D. 20-40 days	
E. Every day	

25. When you smoke how many cigarettes do you usually smoke?

A. 0 cigarettes	
B. 1-4 cigarettes	
C. 5-9 cigarettes	
D. 10-19 cigarettes	
E. 20 or more cigarettes	

The next questions ask about taking drugs (marijuana, cocaine, ecstasy etc.).

26. Have you tried any form of drugs in your lifetime?

A. Yes	1
B. No	0

27. How old were you when you tried drugs, of any form, for the first time?

A. I have never tried any drugs	
B. 10 years old or younger	
C. 11 years old	
D. 12 years old	
E. 13 years old	
F. 14 years old	
G. 15 years old	

28. During the past 12 months, on how many days did you take drugs?

A. 0 days	
B. 1-5 days	
C. 6-20 days	
D. 20-40 days	
E. Every day	

The next questions ask about gangs.

29. Are you a member of a gang?

A. Yes	1
B. No	0

30. Are any of your family members or friends in a gang?

A. Yes	
B. No	

32. Does the gang hurt people?

A. Yes	
B. No	

The next questions ask about things you have done over the past year.

How often in the last year have you …?		Not at all	Just 1 time	2-5 times	6 plus times
33	… travelled on a bus or train without paying enough money or using someone else's pass?	0	1	1	1
34	… written things or sprayed paint on property that did not belong to you?	0	1	1	1
35	… stolen something from a shop or store?	0	1	1	1
36	… sold an illegal drug to someone?	0	1	1	1
37	… ridden in a stolen car or van or on a stolen motorbike?	0	1	1	1
38	… broken into a car or van to try and steal something out of it?	0	1	1	1
39	… done each of these things to someone you know:				
	A. ignored them on purpose or left them out of things.	0	1	1	1
	B. said nasty things, slagged them off or called them names.	0	1	1	1
	C. threatened to hurt them.	0	1	1	1
	D. hit, spat or threw stones at them.	0	1	1	1
	E. got other people to do these things.				
40	… broken into a house or building to try and steal something ?	0	1	1	1
41	… hit, kicked or punched a brother or sister on purpose?	0	1	1	1
42	… hit, kicked or punched someone with the intention of really hurting them?	0	1	1	1
43	… deliberately damaged or destroyed property that did not belong to you?	0	1	1	1
44	…sold something that didn't belong to you or that you knew was stolen?	0	1	1	1
45	… stolen any money or property that someone was holding, carrying or wearing at the time?	0	1	1	1
46	… hit or picked on someone because of their race or skin colour?	0	1	1	1
47	… hurt or injured animals or birds on purpose?	0	1	1	1
48	… set fire or tried to set fire to something on purpose?	0	1	1	1
49	… been rowdy or rude in a public place so that people complained or you got into trouble?	0	1	1	1

Appendix 8: Sample endorsement of risk taking behaviours

Question about risk behaviour	**N endorsed risk behaviour (Percent)**
1. When you ride a bicycle do you wear a helmet?	43 (71.7%)
2. When you rollerblade or ride a skateboard, do you wear a helmet?	14 (23.3%)
3. Do you wear a seat belt when riding in a car?	37 (61.7%)
5. Have you ever carried a weapon, such as a gun, knife, or club?	15 (25%)
7. Have you ever been in a physical fight in which you were hurt and had to be treated by a doctor or nurse?	10 (16.7%)
8. Were you in a fight in the past 12 months?	38 (63.3%)
19. Have you ever had an alcoholic drink (other than a few sips)?	28 (46.7%)
21. Have you ever been sick or felt unwell from drinking alcohol?	5 (8.3%)
22. Have you ever tried smoking, even on or two puffs?	12 (20%)
26. Have you tried any form of drugs in your lifetime?	10 (16.7%)
29. Are you a member of a gang?	1 (1.7%)
33. Used bus/train without paying enough?	27 (45%)
34. Written on property that did not belong to you?	14 (23.3)
35. Stolen from a shop?	8 (13%)
36. Sold an illegal drug to someone?	5 (8.3%)
37. Ridden in a stolen vehicle?	4 (6.7%)
38. Broken into a vehicle to steal something?	2 (3.3%)
39A. Ignored someone you know on purpose?	44 (73.3%)
39B. Said nasty things about someone you know?	40 (66.7%)
39C. Threatened to hurt someone you know?	30 (50%)
39D. Hit, spat at someone you know?	7 (11.7%)
40. Broken into a building to steal something?	3 (50%)
41. Kicked or punched a sibling on purpose?	25 (41.7%)
42. Kicked or punched someone with the intention of really hurting them?	28 (46.7%)
43. Damaged or destroyed property that did not belong to you?	14 (23.3%)
44. Sold something that didn't belong to you?	9 (15%)
45. Stolen property that someone was carrying?	6 (10%)
46. Hit or picked on someone because of their race or skin colour?	1 (1.7%)
47. Hurt or injured animals or birds on purpose?	0 (0%)
48. Set fire to something on purpose?	7 (11.7%)
49. Been rowdy in a public place so that you got into trouble?	18 (30%)

Appendix 9: Supplementary analysis of relationship between executive function composites and risk taking

Table 5_Forced entry regression of Executive Function composite scores predicting real-life risk-taking

Predictor Variable	*B*	*95% BCa CI*	*SE B*	*Beta*	*p*
Constant	16.94	9.34, 24.56	3.78		0.001
Planning	-0.25	-0.92, 0.41	0.33	-0.10	0.46
Working Memory	-0.04	-0.44, 0.48	0.23	-0.03	0.83
Cognitive Inhibition	-0.61	-1.00, -0.22	0.20	-0.35	0.004

Confidence intervals and standard errors based on bias corrected accelerated 1000 bootstrap samples

A supplementary regression analysis was conducted to see whether there was a difference in predictive strength of the individual components of the executive function composite score on real-life risk-taking behaviour. These components were entered into the regression model using the forced entry method, where all three predictors are forced into the model simultaneously. In predicting real-world risk-taking, the model was significant overall $F(3,56) = 3.23$, $p < 0.05$. Cognitive inhibition was found to predict risk-taking behaviour after controlling for both planning and working memory abilities. Table 5 shows the results of this regression analysis. Although the executive function composite predicted risk-taking behaviour, there were differential effects of the individual components and neither working memory nor planning ability uniquely predict real-life risk-taking once cognitive inhibition had been taken into account.

Service Evaluation Project

An investigation of the joint work offered by the National and Specialist Child and Adolescent Obsessive Compulsive Disorder Service: advantages, obstacles and suggested improvements.

Superviser: Dr Amita Jassi

Table of Contents

List of Tables and Figures

Abstract

This study was the first investigation of the joint work provided by the National and Specialist Child and Adolescent Obsessive Compulsive Disorder (NSOCD) Service. The study was conducted in order to ensure that the team continued to make improvements to the quality of this service option. Twenty cases that had received this type of treatment were identified. For these cases NSOCD clinicians, young people, parents and their local clinicians were asked for their views on the advantages and obstacles of the joint work process. Results were analysed for key themes that appear to be important for successful joint working to occur. These include communicating the possibility of joint working to all involved at the earliest possible time, working collaboratively with local clinicians in order to ensure adherence to the treatment model, and ensuring consistent communication takes place throughout treatment. These themes were used to construct a more structured joint working package for the NSOCD team to follow. Some aspects of the joint working process which will need further investigation, including whether this service option provides any additional benefit to the NSOCD team's standard treatment, and a better understanding of the rationale for offering the joint work package to clients.

Introduction

Obsessive Compulsive Disorder (OCD) is characterised by obsessions or compulsions, and most frequently by a combination of both. Obsessions are intrusive thoughts or images that produce fear and worry, while compulsions are repetitive behaviours that a person feels driven to perform. These symptoms cause significant distress to the person, interfere with their day to day functioning and are often time consuming (DSM 5: APA, 2013).

The National Institute of Clinical Excellence (NICE) recommends a stepped care model for the treatment of OCD (NICE, 2005). Cognitive behavioural therapy (CBT) with exposure and response prevention (E/RP), involving the family or carers, is the recommended treatment for OCD. If a young person's OCD does not respond to this treatment, they should be offered a combined treatment of CBT with E/RP and a selective serotonin reuptake inhibitor (SSRI) or clomipramine (NICE, 2005). Young people who have more impaired functioning as a result of their OCD, higher levels of comorbidity, or a poor response to initial treatment, should be referred to specialist teams with greater expertise in treating OCD (NICE, 2005).

The National and Specialist OCD (NSOCD) Service is a Tier 4 Child and Adolescent Mental Health Service (CAMHS) and constitutes one of these teams. It provides specialist assessment and treatment for young people up to 18 years of age across the UK, who have a diagnosis of OCD and related disorders, such as body dysmorphic disorder (BDD). Young people who have a comorbid diagnosis of a developmental disorder, such as high functioning autism spectrum disorders (ASD), are also eligible for this service. Between April 2013 and March 2014, the average severity of the OCD of the clients seen in the clinic was in the moderate to severe range at assessment. The average severity of these clients' OCD at end of the treatment lay in the mild range, indicating that treatment with the NSOCD team leads to clinically significant reductions in symptoms.

The NSOCD team has developed a treatment protocol for OCD which is widely used in the clinic (Turner, 2008). The protocol incorporates psychoeducation about the nature of OCD, anxiety in general and the concept of E/RP. In the early stages of treatment the therapist and client construct a hierarchy of the young person's fears. The majority of therapy is spent supporting the client to engage in ERP tasks that progress up the hierarchy. The final part of treatment is focussed on relapse prevention.

The standard package that the team offers is 14 weekly sessions of CBT. They also provide a package of 20 weekly sessions for those with co-morbidities such as ASD. This is an intensive package which involves 15 hours of therapy over five days, with 4 weekly sessions thereafter and treatment within an adolescent inpatient unit. In addition, the service also offers consultation and joint working, but these treatment packages are less developed than the other service options and there is no formal structure for how they should be offered. This way of working is currently offered case-by-case and it has not yet been evaluated. Such evaluations are important to conduct as "NHS organisations are accountable for continuously improving the quality of their services and safeguarding high standards of care" (Scally & Donaldson, 1998).

Joint working

The NSOCD team often works jointly with clinicians from the client's local CAMHS services. Clinicians from both teams are involved in direct intervention with the client, and also communicate with each other between sessions. Alternatively, the NSOCD team can offer ongoing consultation, where the two therapists communicate about the client's treatment, while the intervention itself is provided by the local therapist alone. To date, the majority of the NSOCD team's clients have received the former type of intervention. Although these interventions are different, they do both involve joint working on the part of the clinicians. Therefore, from this point forwards both will be referred to as 'joint work'.

Every Child Matters (DfES, 2003) and the National Service Framework for Children (DoH, 2004) both highlight the importance of ensuring effective partnerships between all agencies working with a young person. Such partnerships are reported to improve a young person's experience of services and can lead to improved outcomes. These policies emphasise the importance of ensuring effective partnership working when a child has been referred to a specialist CAMHS service, as was the situation for the NSOCD clients who received this type of work. In order to evaluate the strength of these working partnerships, it was necessary to ask for the views of the clinicians working in both the NSOCD team and in the local teams.

The National Service Framework for Children (DoH, 2004) also clearly lays out the importance of seeking the views of service users and incorporating these into reviews of

service provision. Accessing these views is necessary to ensure that service users' needs are at the forefront of service development (Day, 2008; Worrall-Davies & Marino-Francis, 2008). Every Child Matters (2003) reported on the difficulty that young people often face when they are in contact with many professionals who may work in different teams, rather than having one trusted adult who provides continuity of support. In order to access clients' needs and to verify whether the intervention offered so far was meeting these, it was necessary to ask for feedback from the young people who received joint work from the NSOCD team.

Finally, NICE recommends that families or carers should be involved at all stages of a young person's assessment and treatment if appropriate (NICE, 2005). It was therefore also important to ask clients' families for their views on the intervention that their child had received, particularly as it has been found that there can be a low correlation between the satisfaction ratings of young people and of their parents (e.g. Garland, Haine, & Lewczyk Boxmeyer, 2007).

Given the above, this service evaluation set out to ask NSOCD clinicians, local clinicians, clients and their families for their views on the advantages, obstacles and possible improvements of the joint working that was provided by the NSOCD team. It was hoped that the results of this service evaluation study could then be used to propose a more structured model for providing joint work, which would take into account the perspectives and needs of all four groups who are involved in this joint working treatment. This would ensure that improvements would be made to the quality of care and the standard of the joint working received by the NSOCD team's clients.

Aims of evaluation of joint working provided by NSOCD team

1. To investigate the rationale for using joint work from the perspective of the NSOCD clinicians.
2. To ask the NSOCD team and local team for specifics on what was provided.
3. To ask about the advantages, the obstacles and possible improvements from the perspectives of the NSOCD team, the local teams, the clients and their families.
4. To use the findings to suggest a formal structure to follow when the NSOCD team provides joint work.

Method

Participants

Each NSOCD team therapist was asked to identify any cases where they had worked jointly with a local clinician, or offered consultation. Eighteen clients were identified who had received joint working, and two consultation cases were also identified. In total, 20 semi-structured interviews were conducted with the NSOCD team clinicians. The local clinicians who had worked on these cases, as well as the clients themselves and their parents were contacted and asked to complete a questionnaire. Of the 20 local clinicians contacted, 6 responded (30%), of the 20 clients contacted, 6 responded (30%) and of the 38 parents contacted 12 of the parents (32%) responded about 9 clients (a mother and father each completed a separate questionnaire for 3 clients).

All young people who are seen in the clinic complete the Children's Yale-Brown Obsessive Compulsive Scale (CYBOCS). The CYBOCS is a clinician administered interview used to assess symptom severity in young people (Scahill et al., 1997). A CYBOCS score of 30 or above is classified as 'severe' OCD, from 18-30 is 'moderate', and less than 18 is considered 'mild' (Scahill et a., 1997). Of the twenty clients who received their treatment through the joint work model, the average CYBOCS score was 33 at initial assessment, with eighteen scoring 30. All clients who received joint work had a CYBOCS score that was in the upper half of the moderate to severe range on the CYBOCS.

Repeat CYBOCS administration was reviewed for 15 of the 20 clients who received joint work (5 clients had no further CYBOCS scores on their records). The point at which these repeat CYBOCS were completed was very varied, with some completed at the end of treatment, while others were completed at 3, 6 or 12-month follow-ups. The CYBOCS that was administered at the furthest point from the assessment was selected (i.e. the most recent CYBOCS). At this point, when the clients had received at least some treatment through the joint work model, the average CYBOCS score was 18 (mild).

Measures

NSOCD clinician semi-structured interview

A semi-structured interview was conducted with the NSOCD team to allow follow-up questions. The NSOCD clinician's semi-structured interview was designed to investigate why the NSOCD team had provided joint work in the first place. The interview asked for details on what was provided in the joint work and how. Feedback on how it had been to work with the local clinician and advantages, obstacles and potential improvements to the joint working model were also asked for (see Appendix 1).

Local clinician questionnaire

The local clinicians were administered a questionnaire as this took less time than an interview, and so was hoped to elicit a higher number of respondents. The local clinician questionnaire mirrored the NSOCD clinician's interview in that it asked for details about what was actually provided, and for advantages, obstacles and potential improvements. It also asked whether the local clinicians would be interested in receiving further support to build on their therapeutic skills, as this is something the NSOCD team thought they could offer in the future (for example, receiving a manual/training day) (see Appendix 2).

Clients and parents questionnaire

The clients and parents questionnaire was designed to be as concise as possible so as to reduce the time burden. It asked about the advantages, obstacles and potential improvements of receiving support from two teams at the same time, and how beneficial it was to work with both teams (see Appendix 3).

Procedure

The interviews with the NSOCD team clinicians were conducted between November 2012 and April 2013. Local clinicians were then contacted, either by email or telephone between April and November 2013 and asked to complete and return the questionnaire. In July 2013 parents and children were sent a questionnaire to complete by post, and this was followed up until November 2013. Participation in this study was voluntary. Anonymised data was stored electronically in a password-protected database.

Results

Provision of joint work to date.

Why was joint work offered?

For each of their cases, the NSOCD team clinicians were asked why the team had provided joint working (see Table 1). They were offered three options (distance/local clinician already involved/need for ongoing work) and asked for any additional reasons.

Table 1. Why was joint work offered according to the NSOCD team?

	Number of cases (% of total)
Distance	12 (60%)
Local clinician already working with client	11 (55%)
Need for ongoing work	9 (45%)
Additional reasons from NSOCD team interviews	
Complexity	8 (40%)
Needing home based treatment	5 (25%)

The NSOCD clinicians endorsed the three set options at similar rates. Additional reasons included complexity, which was also endorsed at a similar rate, and the need for home based sessions, which was the least likely rationale to be endorsed.

What did the joint work look like?

The majority of local clinicians had psychology roles, while a minority were nurses, primary mental health workers or CBT therapists (see Figure 1).

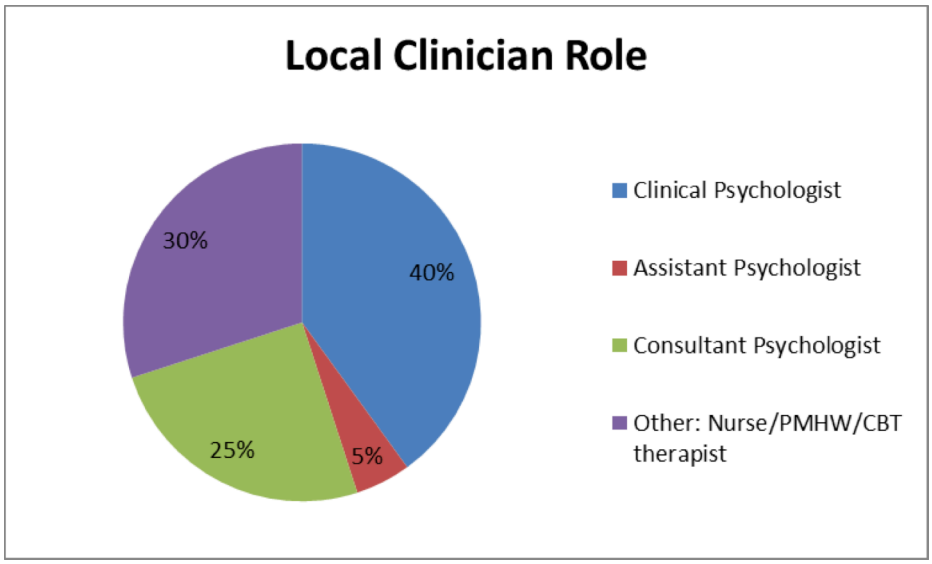


Figure 1_Role of local clinician

The NSOCD clinicians were asked about the content of joint working with local clinicians, when not in direct clinical contact with clients (see Table 2). They were given the options of 'help planning sessions' and 'giving advice', and then asked about any other content. The most common theme was providing help with planning sessions; additional activities provided by the team, such as medication advice and reviews and discussing engagement obstacles, were less frequently endorsed.

Table 2. Content of joint work according to NSOCD clinicians	
	Number of cases (% of total)
Help planning local clinician's sessions	16 (80%)
Giving advice	13 (65%)
Additional themes from NSOCD team interviews	
Broader planning and case management	10 (50%)
Re-formulating e.g. did symptoms relate to comorbid diagnoses.	8 (40%)
Medication advice and reviews	3 (15%)
Discussing engagement obstacles	3 (15%)

Both groups of clinicians were asked about the different ways that joint work was conducted (see Figure 2). It appears that local clinicians felt they had received less observation opportunities than the NSOCD team felt that they had provided, although it may have been that a greater proportion of those local clinicians who did not receive this modality (50% according to the NSOCD team) responded. Similar proportions of NSOCD team and local

team clinicians endorsed the methods of delivering therapy together, and therapy delivery by the local clinician alone.

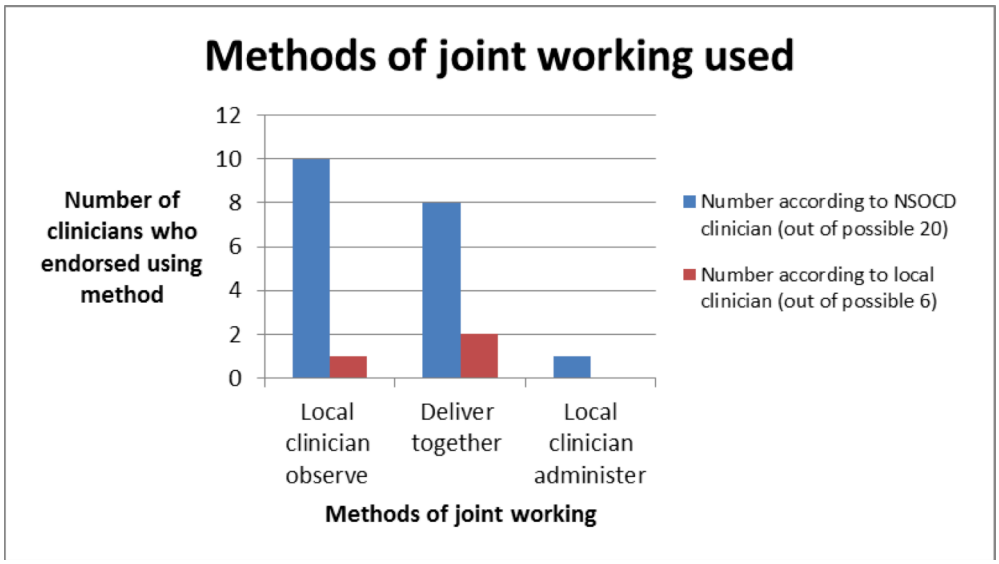


Figure 2_NSOCD and local clinician endorsement of methods of joint working

Both groups of clinicians were also asked how they communicated with each other (see Figure 3). Both groups endorsed communicating over telephone and email more frequently than through face-to-face communication.

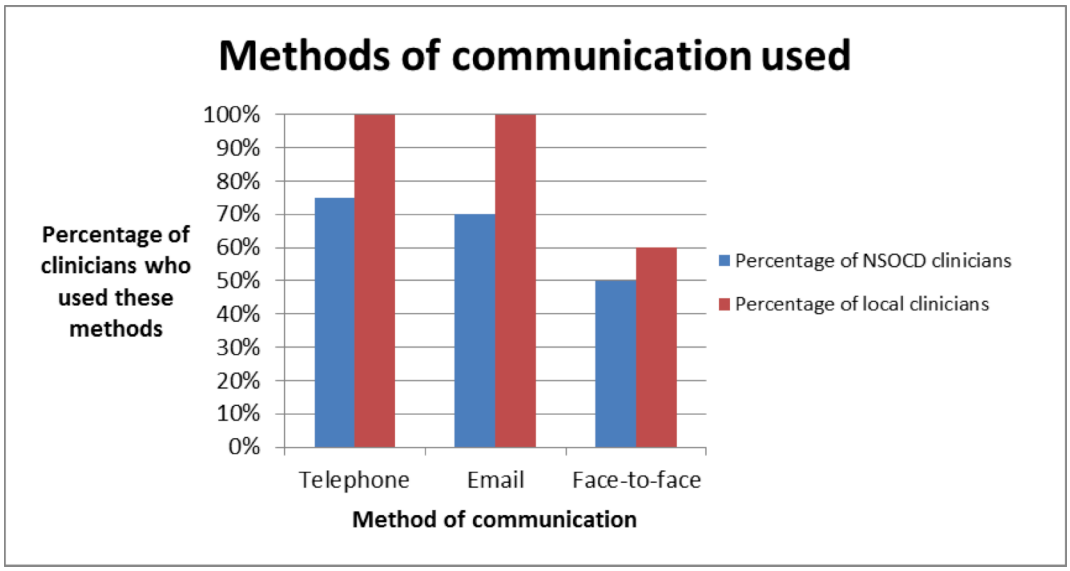


Figure 3_NSOCD and local clinician endorsement of communication methods used

The six local clinicians were evenly split in their preferred method of communication between using the telephone, face-to-face and email (see Figure 4). The NSOCD clinicians

preferred the telephone, followed closely by face-to-face and then email. Many of the reasons given for these preferences concerned a payoff between quality of communication method and feasibility (see Table 3).

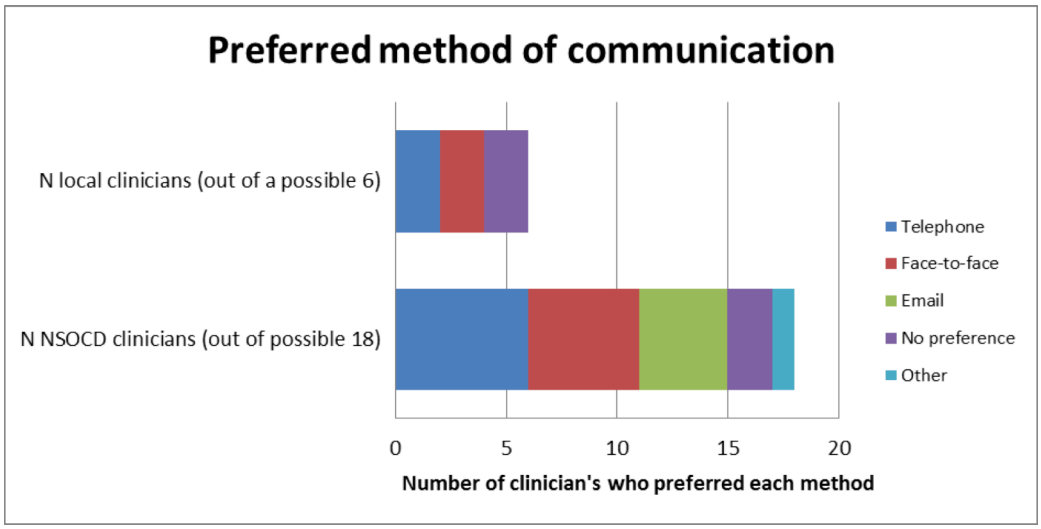


Figure 4_NSOCD and local clinician preference for communication method

Table 3. Why this preference?

	Example quotes NSOCD clinician	*Example quotes local clinician*
Telephone	"feasible but still have 2 way communication"	"two way conversation"
Face-to-face	"modelling E/RP"	"gives observation opportunity"
Email	"written record"	
	"does not require other person"	
Other	"need all methods"	

Was joint work beneficial …

… for clients and their families?

Both clients and parents were asked whether they would have preferred treatment with both teams, or with one of the teams alone (see Figure 5) and then asked to expand on this preference (see Table 4).

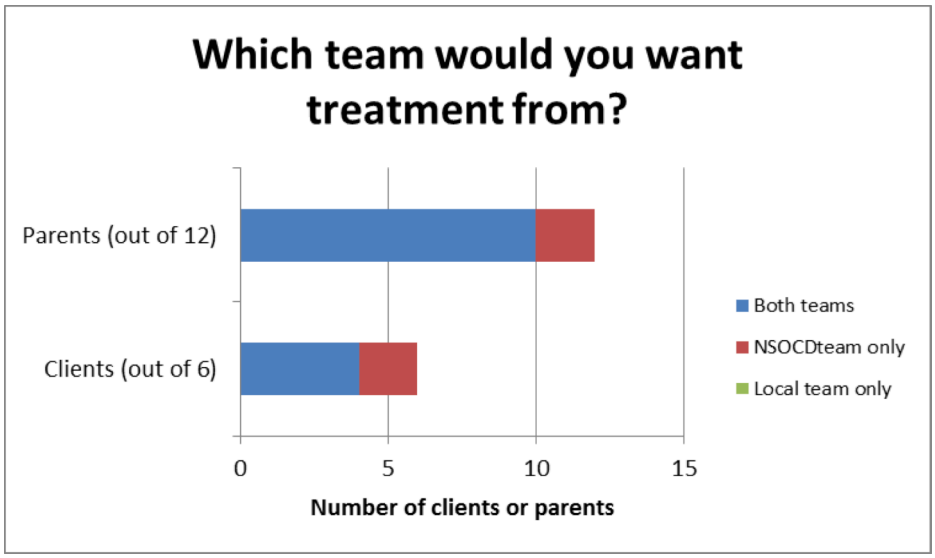


Figure 5_Client and parent preference for treatment provision by team

Figure 5 demonstrates that the majority of both parents and clients preferred treatment with both teams. Reasons for wanting to work with both teams were often related to the understanding that the clients and parents felt they received from the experts in the NSOCD team, which was then augmented by the consistent support the local team could provide. Reasons for wanting to work with the NSOCD team alone were mostly focused on the belief that the NSOCD team provides greater expertise.

Table 4. Why this preference?

	Example quotes client	*Example quotes parents*
Both teams	"NSOCD team knew exactly what I was on about; they really understood what I was trying to explain. The local team were there for support and helped me to practice the tasks I was given."	"NSOCD team were excellent and focused specifically on OCD behaviours. Local clinician supported this work but also able to focus on other trauma in life that impacted on his OCD." "NSOCD team are experts in this field and our child found they knew exactly what she was thinking. Our local team then supported and worked with her on a weekly basis, giving her encouragement"
NSOCD team only	"because they are more experienced"	OCD "team have higher expertise and experience, also more experience of treating OCD with comorbid ASD"

… for local clinicians?

Local clinicians were asked to rate how much they felt the joint work experience had helped them to develop their competency in delivering CBT for OCD on a scale from not changed (0% changed) to greatly improved (100% changed). The five local clinicians who completed this section reported an average change in their competency of 55%. They were also asked to rate how much working with the NSOCD team therapist had developed their confidence in delivering CBT for on a scale from not changed (0% changed) to greatly improved (100% changed). The five local clinicians who completed this section reported an average subjective change in their confidence in delivering CBT of 53%.

What helped the joint working to run well?

The NSOCD team clinicians were asked about local clinician qualities that they thought were beneficial for the joint working (see Table 5 and Appendix 4 for example quotes).

Table 5. Positive comments from NSOCD team about local team clinicians

Theme	*N endorsed (out of 20)*
Good understanding of model	7 (35%)
Reliable and consistent	6 (30%)

All four groups were asked what they thought the advantages of the joint working were (see Table 6 and Appendix 4 for example quotes). A theme that was endorsed by all four was that having two teams increased the capacity to provide a more comprehensive care package. The groups all agreed that the treatment was strengthened when both teams provided consistent messages about the client's treatment. One parent commented that this consistency helped to support their child in fighting against their OCD outside of sessions.

The NSOCD clinicians also reported that the ability to share both the physical workload with the local clinician, and to provide emotional support for each other, enhanced the comprehensiveness of the treatment package the clinicians could provide. NSOCD clinicians and parents reported that there was greater support available when two teams were involved and this was particularly helpful for working with risk.

The NSOCD team also highlighted the importance of the local team's increased proximity for both risk and for working with young people who were house-bound or who required home

based work due to the nature of their OCD. The NSOCD clinicians and local clinicians both mentioned that each team bought unique knowledge to the case. While the NSOCD team was able to offer their expertise regarding OCD itself, local clinicians were able to provide very valuable background information if they had already built up a relationship with the client, which then aided formulation and treatment.

Finally the NSOCD team clinicians discussed the benefit of consistent communication. One clinician mentioned such consistency was supported if clinicians scheduled regular phone calls with each other.

Table 6. Advantages of joint working and what worked well

Theme	*Group*	*Number endorsed*
Two teams can provide more comprehensive care provision	NSOCD clinicians	7 (of 20) 35%
	Local Clinicians	3 (of 6) 50%
	Clients	3 (of 6) 50%
	Parents	5 (of 12) 42%
Greater support can be provided.	NSOCD clinicians	4 (of 20) 20%
	Parents	8 (of 12) 75%
Each team can provide unique information	NSOCD clinicians	2 (of 20) 10%
	Local Clinicians	6 (of 6) 100%
Consistent communication supports effective intervention	NSOCD clinicians	6 (of 20) 30%

How could the joint working process be improved?

The NSOCD team clinicians reported on some features that made it difficult to work with local clinicians and that could impact negatively on the treatment (see Table 7 and Appendix 4 for example quotes).

Table 7. NSOCD team: obstacles to working with local clinician

Theme	*N endorsed (out of 20)*
Inconsistent and disorganised	3 (15%)
Lack of support from local clinician's team	2 (10%)
Not adhering to the model	2 (10%)

All four groups were asked for their views on the obstacles to the joint working (see Table 8 and Appendix 4 for example quotes). Both groups of clinicians said that it was difficult to coordinate contact with each other and that the distance and resulting lack of face-to-face contact raised obstacles. They also said that the clients and their families could present differently with each of the teams. The clients and parents both reported that at times they felt confused as they were getting mixed messages from the two teams.

The NSOCD clinicians said that the work was impeded when the two team's expectations differed. For example one NSOCD clinician reported that the local team appeared to expect that the NSOCD team would carry clinical responsibility alone rather than jointly. One of the parents reported that they felt that the local team was not confident in their ability to manage exposure sessions.

Table 8. Obstacles to joint working

Theme	*Group*	*N endorsed*
Coordinating contact between clinicians	Local Clinicians	4 (of 6) 67%
	NSOCD clinicians	5 (of 20) 25%
Distance	Local Clinicians	2 (of 6) 33%
	NSOCD clinicians	3 (of 20) 15%
Differing expectations of local clinician's role.	NSOCD clinicians	6 (of 20) 33%
Confusing	Clients	2 (of 6) 33%
	Parents	1 (of 12) 8%
Different presentation with each team	NSOCD clinicians	2 (of 20) 10%
	Local Clinicians	1 (of 6) 17%
Concern about local clinician's ability to support E/RP	Parents	1 (of 12) 8%

All groups were asked about areas of the joint working process that they thought could be improved (see Table 9 and Appendix 4 for example quotes). The local clinicians did not provide any responses to this question. The NSOCD clinicians and parents both reported that trying to make the communication between teams as consistent as possible would be of great benefit to the treatment. The NSOCD clinicians thought it would also help to increase face-to-face contact, and that planning in time to talk would be of benefit. Both clients and parents said it would be good to have greater OCD specialisation in their local areas.

Over half of the NSOCD team clinicians said it would help for both teams to agree on what each expected from the other as early as possible. These included logistical expectations

(e.g. local clinician availability), along with ensuring local clinicians were clear about the treatment model that the NSOCD team uses. The NSOCD clinicians also thought it would be beneficial to have more face-to-face sessions, in order to increase direct contact with the client and to provide the local clinician with exposure to running E/RP.

Table 9. Areas to improve

Theme	*Group*	*Number endorsed*
Early agreement on expectations	NSOCD clinicians	12 (of 20) 60%
Consistent, good communication	NSOCD clinicians	5 (of 20) 25%
	Parents	1 (of 12) 8%
Giving joint sessions	NSOCD clinicians	5 (of 20) 25%
Local specialisation	Clients	1 (of 6) 17%
	Parents	1 (of 12) 8%

Local clinicians were asked whether they would have been interested in receiving additional training and support through the joint work process (see Table 10).

Table 10. Local clinicians interest in receiving additional support for clinical work (max = 6)

Training day	3 (50%)
Manual	3 (50%)
Watching recordings	3 (50%)
Group supervision	3 (50%)
NSOCD team observe you	2 (33%)
Record sessions for NSOCD team to listen to	2 (33%)

Discussion

The discussion first considers whether joint working appeared to be of benefit to the clients, their parents and the local clinicians. Following this, the discussion is split into two sections. Section 1 summarises the key themes from the results, which appear to be important for successful joint work, with reference to supporting literature. A potential treatment package for the NSOCD team's joint working, incorporating these themes, is suggested. Section 2 focuses on the feedback of the findings and proposed package to the NSOCD team. The discussions during this feedback session, which focused on the decision-making and the communication process important for setting up the joint working, are reported. The discussion concludes by considering questions about the NSOCD team's joint working that have not been answered by this investigation.

Benefit of joint work

The majority of clients and parents who responded said they would have preferred treatment to be provided by both teams rather than either the NSOCD team or their local team alone. Over the course of treatment, clients experienced a reduction in their symptom severity, with the average CYBOCS score at end of treatment lying in the mild range. This reduction is in line with the trend observed in the overall clinic and it is not possible to know whether joint work provided any added benefit. However the reduction in the severity of client's scores from the severe range to the mild range from receiving therapy through the joint working model shows that clients do experience a significant improvement in their functioning. In conclusion, this model of therapy provision is perceived as beneficial by most of the young people and their families who received care through this format, and does not prevent significant clinical improvements from taking place.

In addition the local clinicians indicated that they felt both their competence and confidence in delivering CBT for OCD was improved through the joint work process. This was a subjective observation but, if valid, these changes are beneficial for both the local clinician and for the service that they are working in and would provide support for the call by some of the parents and clients for increasing local provision in the treatment of OCD.

Section 1: Important factors for successful joint working

The forms that the joint work took with each of the cases investigated here were heterogenous and varied. However through asking the clients, parents, NSOCD clinicians and local clinicians about what worked well and what made things more difficult it was possible to identify several factors that appear to be critical for successful joint work treatment to occur. These factors include the importance of communicating at the earliest possible time with the local team about the possibility of working jointly, of working collaboratively with local clinicians to support adherence to the treatment model and of ensuring consistent communication takes place throughout the joint work.

Communicating with the local team as soon as possible about joint work

The NSOCD team clinicians reported that an obstacle to joint working was their having differing expectations of what was entailed in joint work from the local clinicians. They also reported that joint work was limited when the local clinician's team was not in support of the local clinician's contribution to the joint work. Nearly two thirds of the NSOCD team clinicians said that the joint work would have been improved if there had been an early agreement on what was expected from all involved.

The presence of formal agreements between consultant and consultee organizations, particularly when there is endorsement from the highest levels of the consultee organization, has been found to support the achievement of consultation aims (Caplan, Caplan, & Erchul, 1994). Indeed thorough planning and organization have been found to lead to good outcomes in consultation (Bundy & Poppen, 1986). The research evidence for using joint and integrated working in social care services also indicated that this work benefits when clear legal and financial frameworks are put in place. In addition these partnerships were supported when the involved professionals understood the aims and objectives of the joint work and when professionals' roles and responsibilities were made explicit (Cameron & Lart, 2003). The findings from this literature, along with the feedback from NSOCD team clinicians, indicate that the joint work will be better supported if a discussion of what is involved in this model takes place with the local clinician's team at the earliest possible time point.

What is the benefit of taking part in joint work for the local team?

The psychiatrist Gerald Caplan realised that the mental health of many more clients under his care could be improved if he consulted with other staff members, than if he were to work with clients directly (Caplan et al., 1994). The goals of Caplanian consultation are to improve the consultee's ability to work with the client, and to develop the consultee's competence through the consultation process so that they may work with similar problems independently when encountering them in the future (Orford, 1993).

The joint work provided by the NSOCD team is not exactly based on Caplanian consultation and so its aims are more varied. However if the subjective improvement that local clinicians endorsed in their competence in using CBT for OCD is maintained, then they will have an

improved capacity to support other young people with OCD who present to their local team in the future. This is an important benefit for the local team; in the short term they provide resources (local clinician time) but in the long term they gain greater local competence in OCD treatment. That is, when communicating with the local team about the possibility of joint working, it is not just that such work may help to reduce the psychological distress of the clients under their care, but that the competency and confidence of their clinicians may also be developed through taking part in the work.

Working from a collaborative stance may support adherence to the model

When clinicians work collaboratively …

The research on therapeutic alliance indicates that when a client endorses the tasks of therapy and feels that the therapy is collaborative, therapy outcomes are likely to be more positive (Horvath & Luborsky, 1993). It is likely that the working alliance between both therapists will be strengthened by similar factors. Indeed it has been found that using a collaborative rather than an expert approach during the consultation process increases the level of integrity with which a consultee applies an intervention (Kelleher, Riley-Tillman, & Power, 2008). If the working alliance between clinicians is non-hierarchical and without power differentials, and the clinician who is receiving consultative advice feels free to accept or reject what the other is suggesting, then the hypothesis is that he or she will act upon ideas that make sense in resolving problems (Caplan et al., 1994). It is suggested that achieving a collaborative rather than expert stance in the working alliance between local and NSOCD clinicians will be supported by involving the local clinician in treatment planning from the earliest possible time. Such involvement will preferably begin prior to the start of the client's treatment.

This investigation highlighted the variation in the roles of local clinicians. Although not specifically investigated here, these differing roles are likely to be associated with differing levels of experience and knowledge. This indicates that the amount of support and guidance that an NSOCD team clinician may provide is likely to vary depending on the local clinician they are working with. In order for a local clinician to be able to think about the amount of guidance they would find beneficial, it is important that they are informed about the treatment model used by the NSOCD team. This will allow both clinicians to discuss the local clinician's experience with this model and their confidence in using it, prior to beginning

treatment proper. A study by Evans, Law, Turner, Rogers, and Cohen (2011) investigated care staff's experience of receiving psychological consultation and suggested that consultees would benefit from having information provided before the consultation process began.

… adherence to the treatment model may be increased

The NSOCD team's treatment model focuses on exposure and response prevention. Abramowitz's (1996) meta-analysis reported that this therapy component is effective in treating OCD and found that therapist supervised exposure and complete response prevention were associated with greater symptom improvements (Abramowitz, 1996). Having such a sound theoretical basis is positive, as this factor has been found to improve consultation outcomes (Bardon, 1985).

The NSOCD team clinicians reported that it was important for the local clinicians to have a good understanding of the treatment model used by the NSOCD team, while it was an obstacle if the local clinician did not adhere to the model. This is unsurprising given that adherence to cognitive therapy protocols has been linked to therapy outcomes (Padesky & Salkovskis, 1996).

The OCD team's treatment tool incorporates E/RP as its main therapeutic tool. Some of the NSOCD team clinicians and the parents fed-back about the importance of providing repeated exposure sessions. Local clinicians also suggested they benefitted from support in using this technique and NSOCD clinicians thought it would be beneficial for local clinicians to observe E/RP sessions. This feedback indicates that it would be helpful for local clinicians to be included in the initial sessions, when the rationale for the treatment modality is set up for clients and their families.

A further predictor of good consultative outcomes is when both groups agree on what the problem is (Kurpius, Fuqua, & Rozecki, 1993). These initial treatment sessions involve the development of a collaborative formulation of the client's OCD during which the clients and clinicians (and parents) come to a mutual understanding of the problem. It is suggested that the local clinician's presence in these sessions is very important to ensure that all groups involved in the joint work have an agreed understanding of the problem. This mutual

understanding will form the foundation of, and rationale for, adherence to the treatment model in the sessions to follow.

Consistent communication is important

All four groups indicated that consistent and good communication was an important factor in facilitating good joint working.

Poor communication is an obstacle to the joint work

The greatest obstacle to receiving care through this format reported by clients and their families was that it could be confusing at times. A small number of local and NSOCD team clinicians reported that joint work was made difficult when clients presented differently to each team's clinicians. Reliable and regular communication between clinician's would highlight differences. A quarter of the NSOCD team clinicians and a few parents suggested that achieving consistent communication was an important improvement to achieve.

Good communication supports the joint work

On the other hand, approximately one third of the NSOCD team clinicians said that good communication led to good joint work treatment. One outcome of consistent communication is that clients and their families would receive a coherent message and action plan from both therapists. Consistency and the ability to maintain the therapeutic framework have been found to be important therapist attributes for building strong therapeutic alliances early in the relationship (Horvath & Luborsky, 1993). A meta-analysis of the impact of therapeutic alliance found that it was a consistent predictor of treatment outcome (Horvath, Del Re, Flückiger, & Symonds, 2011). Although the literature on therapeutic alliance has focused on single therapist-client dyads, it is likely that consistency will be just as important to achieve (if not more so) if treatment is provided by more than one clinician.

Consistent communication also seems to be supportive for clinicians themselves. In specific quotes (see Appendix 4, Table 6) from clinicians, communication between clinicians was a process by which clinicians could find support if the complexity of a case was overwhelming. In addition, when commenting on the unique support that each clinician could bring to the

joint work model, some local clinician's comments included the benefit of receiving reassurance around using E/RP.

Good communication needs support

Approximately one third of NSOCD team clinicians and all of the local clinicians reported that they communicated with each other via the telephone and email, but only approximately half of them reported using face-to-face communication. This is despite clinicians commenting on the importance of having face-to-face communication. Understandably this form of communication was most difficult to organize due to the need for clinicians to travel and be available for a longer period. Therefore it appears that this aspect of communication when joint working may need more support (in terms of time and resources) than it is currently receiving. Clinicians also reported that telephone conversations were less feasible than email contact, although the benefit of having two-way communication was reported. For example, one NSOCD team clinician said that it would have been helpful to discuss treatment without having the family present. It appears that this form of communication may also require further support.

Two thirds of local clinicians and a quarter of NSOCD team clinicians reported that contact between each other was difficult to coordinate. The importance of having well organized contact can be seen through feedback from the NSOCD team clinicians; when local clinicians were reliable and consistent the working relationship was improved, while it was more difficult to work with inconsistent and disorganized local clinicians. One way in which to increase support for such contact is to set expectations as to the ideal amount of contact as early as possible with clinicians, their teams and the clients and their families. It would may be beneficial to set up dates and times when communication (and what type of communication) will occur in order to reserve time and space in busy clinician's diaries.

In conclusion it is suggested that communicating with the local team about the feasibility of joint work should take place at the earliest possible point and, if possible, a formal agreement on what is expected from each team should be drawn up prior to beginning treatment with the client. It is important for local clinicians to know which treatment model the NSOCD team uses and the rational for this to be communicated at the earliest opportunity. It is hoped that this can support the development of a collaborative

relationship between the clinicians, which in turn can support adherence to the treatment model. Consistent communication between clinicians is also important to achieve. Those involved in the joint working investigated here have suggested that such communication can support the likelihood that clients and their families receive a coherent message and action plan. They have also indicated that through good communication the clinicians themselves can offer support and reassurance to each other.

Proposal of a joint working package

The proposed joint work package (see Figure 6) is thought to incorporate the factors discussed above. This package involves two stages: an initial planning stage, followed by the delivery of treatment. The suggested treatment protocol is based on the NSOCD team's general treatment protocol (Turner, 2008).

Once the NSOCD team has decided that a client may benefit from joint work (see below), the joint work enters the planning stage. At this point this possibility needs to be communicated with the client's local team. If the local team agrees to support such work then a conversation between the OCD and local clinician needs to take place during which they can discuss the structure of the joint work (for example, how much communication they would like to have outside the direct client contact). Following this a formal agreement can be drawn up between the clinicians and their teams on what is expected from all sides during the joint work. The second stage involves the provision of treatment to the client through the joint work model. This incorporates both direct clinical contact and indirect communication between the clinicians.

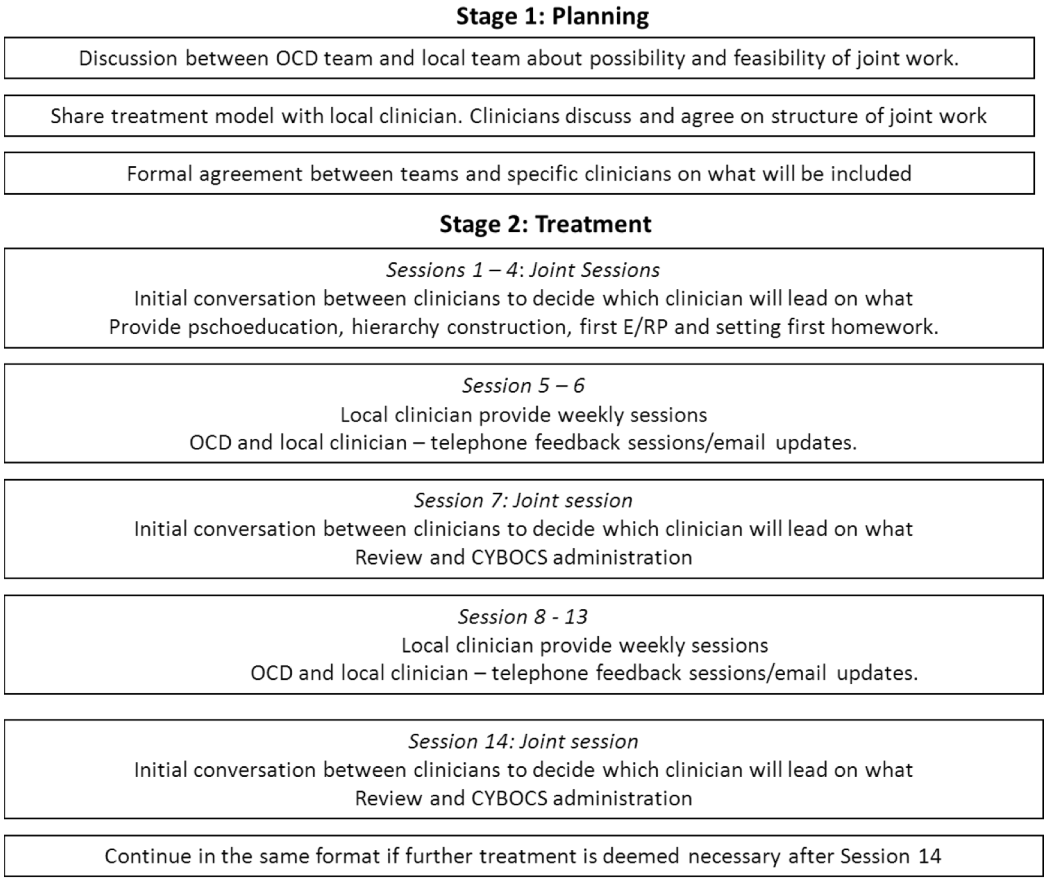


Figure 6_Suggested package for the NSOCD teams joint working

Section 2: Feedback to the service

These results and suggested package were presented to the NSOCD team. This feedback session included both the clinical psychologists who had already taken part in the interviews, along with psychiatrists and other clinicians who had not been involved previously. The feedback focused on asking the team for their opinions on the decision making and communication processes necessary for successful setting up of joint working.

Feedback on the package

The package received a positive response from the team, with clinicians reporting that they were beginning to loosely follow such a structure. Clinicians said it is important, but often difficult to achieve a collaborative (not expert) stance within the joint working relationship. This corresponds with reports from consultants who also felt that maintaining such a

collaborative stance with consultees during a consultation process is a significant challenge (Erchul, 1993). The NSOCD team clinicians felt that an initial face-to-face discussion between both clinicians, without the client or their family being present, was particularly important to ensure a collaborative clinical team was set up. It was also suggested that it would be important to plan to have joint sessions (i.e. sessions 1-4, session 7 etc) together, and to decide who would lead on which sections of the session. This would require clinicians to meet together before direct work begins with the client.

Predicting which clients will need joint work

The earlier the team is able to decide which clients need joint work, the earlier they will be able to communicate the possibility of working in this format to a local team. According to the interviews with the NSOCD clinicians, distance, the involvement of a local clinician, the need for ongoing work, complexity of the client's presentation and the need for home based work, were all factors that indicated joint work would be needed. However these factors were not endorsed for every client that received joint work.

During the feedback session, the NSOCD team did not feel that distance was an important factor indicating the need for joint work. For example, sometimes clients and their parents are prepared to travel long distances to the NSOCD team's clinic. In addition the NSOCD team also provides treatment over the telephone and so distance itself does not require the addition of the local clinician into the joint work model. The team felt that the degree of complexity in a client's presentation was a greater indicator of the appropriateness of joint working. Given that the clients who received joint work scored higher on the CYBOCS when compared to the clinic average, it may be that severity of OCD is an important predicting factor. The team concluded that they do not feel that they are at a point to systematically predict which clients will need this type of work. However there was some agreement that the decision about which clients would benefit from joint work could be made at assessment.

When to communicate about the possibility of joint work?

This investigation highlighted the importance of having conversations about the possibility of the need for joint work at the earliest possible time. During feedback to the NSOCD team, it was suggested that if this model seems suitable, then families could be informed about

the possibility of receiving treatment through this model at the end of the assessment. At this point they could be asked about how much support they are currently receiving from the local team and whether they would be interested in using this model.

There was some discussion during the feedback session about the possibility of including the referrer during the assessment process but the team was split as to whether this was a good idea or not. Some commented that it is important for a client and their family to feel free to tell their story including their experience of previous treatment. Others felt that it could be important to hear from the local team as to what treatment they felt had been provided. If the local clinician was invited to the assessment, this could provide a useful foundation upon which to set up joint working.

Decisions still to be made

The overall conclusion from the NSOCD team was that they were not yet at a stage to formalise the decision-making and communication process around setting up the joint work. Several questions about this process, which were not answered by the initial data, were discussed during the feedback session, but the team did not yet feel they were at a point to decide on answers. These questions are presented below.

1. Which clients need joint work?

   Although this investigation has not elicited factors that could allow the team to predict which clients will benefit from this model with confidence, it does seem that the severity of a client's presentation is an indicator. In addition there was some indication that at assessment, clinicians often 'feel' that the joint work model may be useful.

2. When should the local team be approached about the possibility of joint work?

   A tentative suggestion offered during the NSOCD team's feedback to these results, was that this could take place once clients and parents have given their agreement to working in this model at assessment.

3. Which member of the NSOCD team should contact the local team?

   Ideally it would be a clinician who was involved with the assessment process.

4. What needs to take place during the conversation with the local team?

   It will be important to communicate the content of the joint work package and why it is beneficial. It will also be necessary to find out what can realistically be provided by the local team (e.g. clinician availability, length of time available for). It may be that the local team will need time to discuss this before committing to the model. In this case, which clinician in the NSOCD team will follow this up?

Limitations

The first aim of this study was to investigate the rationale for providing joint work. This research has indicated that there are many different reasons. For example there is some indication that the rationale may be related to the complexity of client's presentations. However this project did not elicit a consistent rationale. During the interviews the NSOCD clinicians found it hard to recall clearly the reasons for providing joint work. It is hoped that going forward, the rationale for providing joint work can be documented at the time when the decision is made, which may help to define more firmly the rationale for joint work.

It is important to note that there was a lack of feedback from local clinicians and clients, in comparison to parents and NSOCD clinicians. The views of the former two groups are therefore underrepresented here, limiting confidence in the extent to which the responses which were received, are representative. In order to increase response rates it may be beneficial to ask for feedback from clients, parents and local clinicians, directly on completion of the joint work. The questionnaires designed for this study (see Appendices 2 and 3) could be modified to investigate other questions of interest that have arisen from this investigation for example.

The measures used in this study were developed specifically for investigating the NSOCD team's joint work and therefore they are not standardized. This was due to the exploratory nature of this initial evaluation of the joint work process. It is hoped that this investigation will allow the NSOCD team to begin standardising the joint work that they provide, and therefore further investigations into the NSOCD team's joint work may be able to administer

more objective measures. In addition, this investigation did not use a formal qualitative methodology for conducting the thematic analysis.

Finally, it is not clear from this investigation whether joint work provides any additional benefit compared with standard treatment with the NSOCD clinician alone. In the future, once the proposed model has been used routinely, it may be useful to compare treatment progress and outcomes in those who do and do not receive joint work. It is hoped that as a consistent joint working process is developed by the team, it may be possible to investigate the impact of joint work in comparison to matched controls. In addition it is not clear that the joint work increases the skills of the local clinicians. Although this investigation reported local clinicians' subjective experience of improvements in their own confidence and competence, there may be other, more objective means that could be used to investigate whether the joint work provided by the NSOCD team really does increase the skills of mental health professionals in assessing and treating OCD (NICE, 2005).

Future directions

This investigation has highlighted several areas that for further investigation. Firstly it is recommended that the team review the joint work package (Figure 6) after following it for the next ten joint working cases. It will be important to review the efficiency of the joint work model suggested here in order to ensure continuing improvements in the quality of care that the OCD clients are receiving (Every Child Matters, 2003). For example, the NSOCD team clinicians were asked in this investigation about the content of the conversations they held with the local clinicians during the joint work process. It may be helpful to present these themes to local clinicians to find out whether there are particular aspects that are more helpful than others.

It will also be very helpful to be able to predict with more confidence, which clients may benefit from joint work. It is suggested that systematic recording of possible factors should now be put in place, so that investigations to identify which are the important factors can be conducted in the future. Examples of factors include those reported in Table 1 of the results (for example, complexity and requirement for home-based treatment). It is also recommended that the rationale for providing joint work be clearly documented.

Finally it would be interesting to investigate whether there is a difference in outcome between clients who receive joint work (client works with clinicians from two teams) compared to consultation (client works with local clinician only), or when they receive joint work (more generally) compared to telephone treatment from the NSOCD team. If these investigations indicate that different modalities lead to different outcomes, then are there factors that indicate a client should receive one particular treatment modality?

Dissemination

The findings from this service project were fed back to the NSOCD team at their away day. The suggested joint working package (Figure 6) will now be sent out to the referrer along with the assessment report if joint work is indicated as an option. The NSOCD team is also intending to add joint working as one of the treatment options that it provides into its service description booklets.

A 'thank you' letter that detailed the main findings of this investigation was also sent out to the clients, parents and local clinicians who participated in the project (see Appendix 5).

Leadership

The National Leadership Council has suggested a framework to promote leadership across the NHS (NLC, 2010), which informs the following discussion of the leadership skills that I used and developed during this service project. Throughout the entire process of conducting this evaluation I have had to manage my time effectively and engage in continuous personal development.

Initially I worked together with two of the NSOCD clinicians to develop the direction of the service, through setting the aims of the investigation and designing the interview and questionnaires. I then launched the project to the rest of the psychology team during a psychology meeting by describing the rationale for the study and the importance of ensuring that the team continued to improve their joint working provision. This presentation was effective in encouraging the team's contribution as I recruited the NSOCD clinician's to complete interviews on all 20 of the joint working cases (100% response rate).

This report has shown how the results of this investigation were used to inform a more structured model of joint working for the NSOCD team. On launching this model to the team during the feedback of the results, I again had to clearly describe the rationale for incorporating this innovation into their everyday work. It appears that this was successfully communicated as the team is now following the model (see Figure 6) with their new joint working cases.

Conclusion

One of the factors that the World Health Organisation (WHO) suggests is important to review when looking at clinical governance is the extent to which patients are satisfied with the service they have received (Every Child Matters, 2003). The clients and parents who responded in this investigation have indicated that they were satisfied with receiving care through this joint work model. The subjective accounts of parents, local clinicians and NSOCD clinicians is that the joint work model increases the support available for managing risk, which is another aspect that the WHO states is important to review during clinical governance. Professional performance is a third factor, and although not objectively measured here, subjectively local clinicians report that their competence and confidence in treating OCD with CBT were improved through this joint work experience. According to NICE, specialist NSOCD teams are expected to increase the skills of mental health professionals in both assessment and treatment of OCD (NICE, 2005), and it appears that joint work may be a way of achieving this expectation. The final aspect of clinical governance that the WHO indicates is important to investigate is resource use, or efficiency. It is hoped that now a joint work model has been suggested, this can be continuously reviewed to ensure resources are used as efficiently as possible.

References

Abramowitz, J. S. (1996). Variants of exposure and response prevention in the treatment of obsessive-compulsive disorder: A meta-analysis. *Behavior Therapy, 27*(4), 583-600.

APA. (2013). *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*: Amer Psychiatric Pub Incorporated.

Bardon, J. I. (1985). On the Verge of a Breakthrough. *The Counseling Psychologist, 13*(3), 355-362.

Bundy, M. L., & Poppen, W. A. (1986). School counselors' effectiveness as consultants: A research review. *Elementary School Guidance & Counseling*.

Cameron, A., & Lart, R. (2003). Factors promoting and obstacles hindering joint working: a systematic review of the research evidence. *Journal of Integrated Care, 11*(2), 9-17.

Caplan, G., Caplan, R. B., & Erchul, W. P. (1994). Caplanian mental health consultation: Historical background and current status. *Consulting Psychology Journal: Practice and Research, 46*(4), 2.

Day, C. (2008). Children's and young people's involvement and participation in mental health care. *Child and adolescent mental health, 13*(1), 2-8.

Department for Education and Skills (DfES). (2003). Every child matters. Norwich:TSO.

Department of Health (DoH). (2004). National service framework for children, young people and maternity services: The mental health and psychological wellbeing of children and young people. London: DoH.

Erchul, W. P. (1993). Selected interpersonal perspectives in consultation research. *School Psychology Quarterly, 8*(1), 38.

Evans, K., Law, H., Turner, R. E., Rogers, A., & Cohen, K. (2011). A Pilot Study Evaluating Care Staffs' Perceptions of their Experience of Psychological Consultation within a Mental Health Setting. *Child Care in Practice, 17*(2), 205-219.

Garland, A. F., Haine, R. A., & Lewczyk Boxmeyer, C. (2007). Determinates of youth and parent satisfaction in usual care psychotherapy. *Evaluation and Program Planning, 30*(1), 45-54.

Horvath, A. O., Del Re, A., Flückiger, C., & Symonds, D. (2011). Alliance in individual psychotherapy. *Psychotherapy, 48*(1), 9.

Horvath, A. O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of consulting and clinical psychology, 61*(4), 561.

Kelleher, C., Riley-Tillman, T. C., & Power, T. J. (2008). An initial comparison of collaborative and expert-driven consultation on treatment integrity. *Journal of Educational and Psychological Consultation, 18*(4), 294-324.

Kurpius, D. J., Fuqua, D. R., & Rozecki, T. (1993). The consulting process: A multidimensional approach. *Journal of Counseling & Development, 71*(6), 601-606.

National Institute of Clinical Excellence (NICE). (2005). Obsessive-compulsive disorder: Core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder.

National Leadership Council (NLC). (2010). Clinical Leadership Competency Framework Project: Report on Findings. *NHS Institute for Innovation and Improvement*.

Orford, J. (1993). *Community psychology: Theory and practice*. Chichester, UK: John Wiley and Sons.

Padesky, C. A., & Salkovskis, P. (1996). Developing cognitive therapist competency: Teaching and supervision models. *Frontiers of cognitive therapy*, 266-292.

Scahill, L., Riddle, M. A., McSwiggin-Hardin, M., Ort, S. I., King, R. A., Goodman, W. K., . . . Leckman, J. F. (1997). Children's Yale-Brown obsessive compulsive scale: reliability and validity. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(6), 844-852.

Scally, G., & Donaldson, L. J. (1998). Clinical governance and the drive for quality improvement in the new NHS in England. *bmj, 317*(7150), 61-65.

Turner, C. (2008). *Learning about OCD and fighting back: A cognitive-behavioural treatment protocol for young people.* London: Obsessive Compulsive and Related Disorders Clinic, Maudsley Hospital, Children's Department.

Worrall-Davies, A., & Marino-Francis, F. (2008). Eliciting children's and young people's views of child and adolescent mental health services: a systematic review of best practice. *Child and adolescent mental health, 13*(1), 9-15.

Appendices

Appendix 1: Interview for OCD team

NS-OCD Therapist: _____

Local Therapist: _____

Patient: _____

1. a. Why did we offer consultation or joint working? Tick all that apply.

☐ Distance

☐ Need for ongoing work after OCD team finished work

☐ Local clinician already working with client

☐ Other _____

b. i. Was there a specific package given?

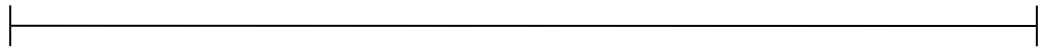
☐ No

☐ Yes – give details below:

---

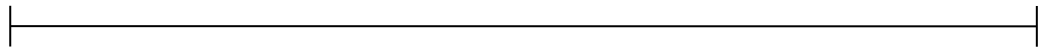
2. How many sessions included the following aspects of consultation/joint work and how helpful did you find them?

a. Having the local therapist observe you give a session. Number of sessions ____



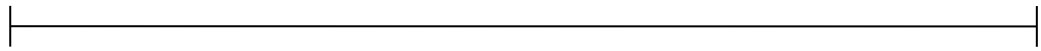
Not at all helpful — Very helpful

b. Observing the local therapist give sessions? Number of sessions ____



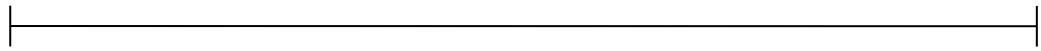
Not at all helpful — Very helpful

c. Delivering sessions together with the local therapist? Number of sessions ____



Not at all helpful — Very helpful

d (i). Outside of session, communication with the local therapist about what either of you did in sessions? Number of times ____. Please tick if you cannot remember number ☐



Not at all helpful — Very helpful

d (ii). During these conversations what did you offer the local therapist? Tick all that apply:

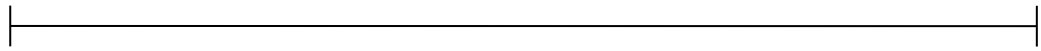
☐ Help with planning sessions

☐ Giving advice

☐ Other _____

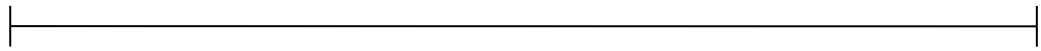
3. How helpful do you think each of these modalities was for communicating with the local therapist, if you used the modality?

a. Face-to-face. Please tick if you did not use this modality ☐



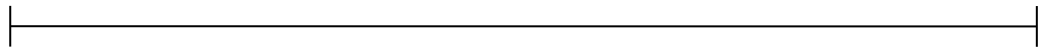
Not at all helpful — Very helpful

b. Telephone. Please tick if you did not use this modality ☐



Not at all helpful — Very helpful

c. Email. Please tick if you did not use this modality ☐



Not at all helpful — Very helpful

d. Did you have a preference for a particular communication modality?

☐ No preference

☐ Face-to-face

☐ Telephone

☐ Email

☐ Other _____

Why this preference?

_____

_____

4. Overall how satisfied were you with:

a. Communication with the local therapist throughout OCD team treatment?

Comments _____

b. Communication with the local therapist through-out OCD team follow-up?

Comments _____

c. Any other comments about consultee?

5. How much do you think that working together with the local therapist enhanced treatment more than either team working with the client separately?

6. a. What were the advantages of joint working with the local therapist?

b. What were obstacles with joint working?

c. Any areas that you now feel could be improved on?

d. Any other comments?

Appendix 2: Interview for local therapist

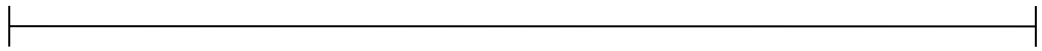
Local Therapist: _____

NS OCD Therapist: _____

Patient: _____

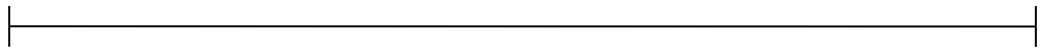
1. How many sessions included the following aspects of consultation/joint work and how helpful did you find it?

a. Observing the OCD team therapist giving sessions. Number of sessions ____



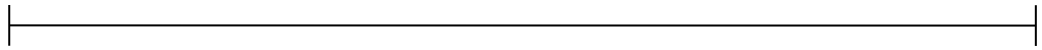
Not at all helpful — Very helpful

b. Having the OCD team therapist observe you giving sessions? Number of sessions ____



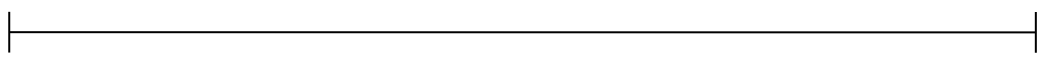
Not at all helpful — Very helpful

c. Delivering sessions together with the OCD team therapist? Number of sessions ____



Not at all helpful — Very helpful

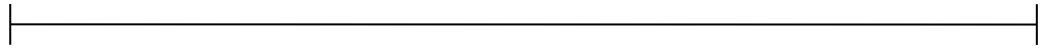
d. Outside of session, communication with the OCD team therapist about what you did in sessions? Number of times ____ .



Not at all helpful — Very helpful

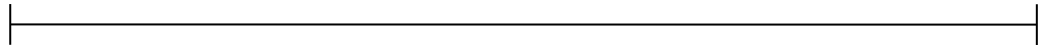
2. How helpful do you think each of these modalities was for communicating with the OCD team therapist, if you used the modality?

a. Face-to-face. Please tick if you did not receive advice in this modality ☐



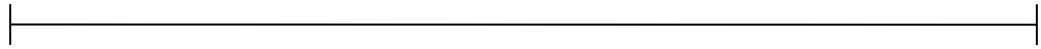
Not at all helpful — Very helpful

b. Telephone. Please tick if you did not receive advice in this modality ☐



Not at all helpful — Very helpful

c. Email. Please tick if you did not receive advice in this modality ☐



Not at all helpful — Very helpful

d. What was the most helpful way of communicating?

☐ No preference

☐ Face-to-face

☐ Telephone

☐ Email

☐ Other _____

Why this preference?

_____

_____

3. Overall how satisfied were you with:

a. Communication with the OCD team therapist throughout OCD team treatment?

Comments _____

b. Communication with the OCD team therapist through-out OCD team follow-up?

Comments _____

4. How much do you think that working together with the OCD team therapist enhanced treatment more than either team working with the client separately?

5. a. What were the advantages of joint working with the OCD team therapist?

_____

_____

b. What were obstacles with joint working?

_____

_____

c. Would any of the following have been helpful and feasible?

☐ Attending a training day beforehand

☐ Having the OCD team observe your sessions

☐ Being given a specific manual protocol to follow

☐ Watching recorded sessions of other cases

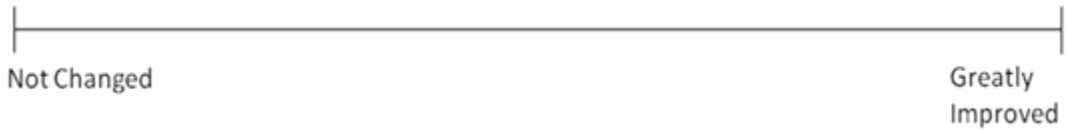
☐ Bringing recordings of your own sessions for the OCD team therapist to comment on

☐ Group case supervision

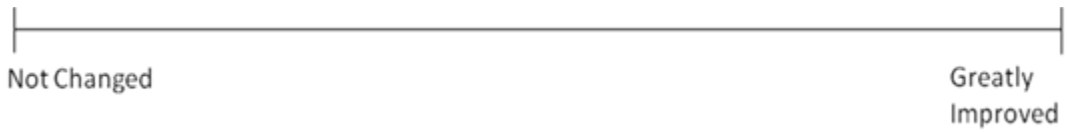
☐ Other ideas

_____

6. a. Do you feel as though working with the OCD team therapist has helped you to develop your competency in delivering CBT for OCD?



b. Do you feel as though working with the OCD team therapist has helped you to develop your confidence in delivering CBT for OCD?



7. Any other comments?

Appendix 3: Interview for clients and families.

Your name: _____

NSOCD clinician: _____

Local Team Clinician: _____

1. Would you have preferred to work with the NSOCD team only, your local team only or both teams?

☐ NSOCD team only

☐ Local team only

☐ Both teams

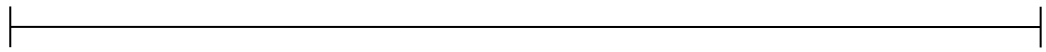
☐ No preference

Why?

_____

_____

2. How satisfied were you with the package of care provided?



Very unsatisfied — Very satisfied

3. What were the benefits of working with both the NSOCD team and a local team?

_____

_____

_____

4. What were the obstacles, e.g. was it confusing to have lots of different people involved?

5. Any areas that you now feel could be improved on?

6. Any other comments?

Appendix 4: Example quotes

Table 5. Positive comments from NSOCD team about local team clinicians

Theme	*Example quotes*
Good understanding of model	"really understood principles"
	"came from psychodynamic background, but picked up on CBT model well."
Reliable and consistent	"stuck to the plan and followed it through."
	"consistent on feedback of her sessions; it was helpful to have this knowledge."
	"implemented what was discussed."

Table 6. Advantages to joint working and what worked well

Group	*Example quotes*
Theme 1: Two teams - more comprehensive care possible	
NSOCD team	"we could provide support for each other with such a complex case". "the local team held general care". "... two clinicians could take on different types of work and provide different aspects of care" "able to provide regular support ... especially important for risk" "parents had reassurance that they were doing the right thing".
Local Clinicians	"meant that there was some treatment continuity, but benefit for client was that most was delivered locally". "continuing therapy between Maudsley sessions."
Clients	"I was supported all the way through by both teams" "the care was continued in the same format"
Parents	"having continuous care and action plan followed throughout programme" "there was an expectation on her to keep the momentum going and focus on fighting OCD in between her London visits".
Theme 2: Capacity for greater support	
NSOCD team	"local support allowed frequent home based work ... this was necessary due to specificity of OCD". "in better position to act if there was risk"
Parents	"Felt more supported and reassured" "we always felt like we had lots of support" "having more people to help in a crisis" "good to have someone local to make contact with and not travel so far" "Feel our daughter got better more quickly through the support we were given from both teams"
Theme 3: Unique support provided by each team	
NSOCD team	"local clinician ... could provide background information and detail".
Local Clinicians	"specialist advice and support and reassurance about using E/RP" "sharing of ideas and joint expertise"
Theme 4: Benefit of consistent communication	
NSOCD team	"scheduled regular phonecalls." "family knew we had contact with the local team and so they knew the information was consistent". "important intervention was for both teams to communicate a united and consistent message".

Table 7. NSOCD team: obstacles to working with local clinician

Theme	*Example quotes*
Inconsistent and disorganised	"did not follow through with everything".
Lack of support from local clinician's team	"her supervisor prevented her from doing E/RP due to health and safety concerns" "remit of team was to assess and hold rather than to treat."
Not adhering to the model	"focused more on cognitive aspects than on E/RP". "did not adhere to the protocol or the model"

Table 8. Obstacles to joint working

Group	*Example quotes*
Theme 1: Difficulty coordinating contact	
NSOCD team	"difficult to coordinate contact with each other's very busy schedules". "... would have been good for us to think through some decisions without the family"
Local Clinicians	"finding time to talk" "matching diaries
Theme 2: Distance	
NSOCD team	"would have been good to have more contact"
Local Clinicians	"... distance ..."
Theme 3: Different presentation with different teams	
NSOCD team	"two teams were getting different messages from client's mother ... important to communicate". "so many people were involved ... difficult to ensure effective communication with everyone"
Local Clinicians	"client displayed different behaviour at different teams"
Theme 4: Confusing	
Clients	"sometimes there were a few mixed messages" "it was confusing to have lots of people involved"
Parents	"was a little confusing at first"
Theme 5: Differing expectations	
NSOCD team	"local team became less involved as soon as I started working" "unclear who was main point of contact during transition from working mainly with NSOCD team to mainly with local team."
Theme 6: Poor confidence in local team	
Parents	"local team unclear about their ability to manage the exposure sessions"

Table 9. Areas to improve

Group	*Example quotes*
Theme 1: Consistent, good communication	
NSOCD team	"more contact outside of direct contact with clients " "should have also met face-to-face" "... good to have set-up times"
Parents	"communication between teams".
Theme 2: Local specialisation	
Clients	"would like more local OCD specialist centres"
Parents	"Think there may be a training need locally - local team seemed to look for distraction techniques rather than facing the fear as OCD therapy does."
Theme 3: Early agreement on expectations	
NSOCD team	"at some point the local team needs to commit to something" "identifying on key person to liaise with" "good to have set up times" "good to be clear about importance of pushing up the hierarchy". "good to find out remit of service before starting treatment and how long local clinician will be available for". "being clear about exposure expectations from very beginning". "need to be very transparent about what we do"
Theme 4: Giving joint sessions	
NSOCD team	"good to do initial face-to-face session." "having more direct contact would have aided local clinicians ... treatment ... good to have direct observation of E/RP" "Would have been helpful to meet the family or be present at the assessment ... this client was actually on ASD spectrum"

Appendix 5: Thank You Letter

Dear all,

We are in the process of writing up the findings from our investigation into whether we can improve the joint work treatment that clients receive from both the OCD team and their local team. We wanted to write to thank you very much for the time you gave to complete our questionnaires about your experience of this work.

It was very positive to see that the majority of clients and their parents who received treatment through this format said they would have wanted treatment to be provided by both teams rather than either the NSOCD team or their local team alone.

You also provided us with constructive feedback. Clients and their parents said that sometimes this method of working could be confusing. Clinician's from local teams wrote about how important it was to make sure the client's experience was coherent. This feedback from you has led us to think very hard about ways that we can try and ensure that communication is as consistent as possible throughout the joint work experience.

Your feedback has been extremely informative and useful, and has led us to think very carefully about how we can improve the quality of the joint work that this team provides with local teams.

With thanks and best wishes,

The OCD team